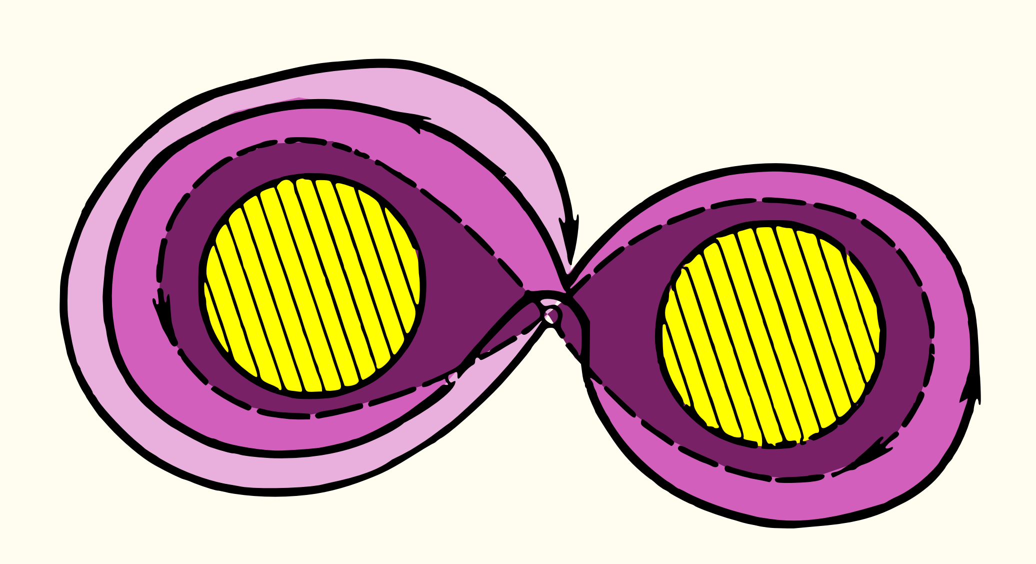
A. Kostrikin

INTRODUCTION Á L'ALGÉBRE



Éditions Mir Moscou

A. KOSTRIKIN

INTRODUCTION À L'ALGÈBRE

ÉDITIONS MIR · MOSCOU

Traduit du russe par
V. KOLIMÉEV

На французском языке

TABLE DES MATIÈRES

AVANT-PROPOS

Le but de ce livre est de donner un exposé systématique du cours d'algèbre tel qu'il s'est réellement établi ces dernières années et professé à la faculté de Mécanique et de Mathématiques de l'Université de Moscou. Une évolution bien naturelle des programmes standard a prédéterminé une modernisation, fût-elle partielle, des manuels d'algèbre. Malheureusement, lors de l'exposé par écrit, le contenu des conférences, enrichi de nombreux détails, s'est fortement déformé.

Formellement, le livre est divisé en deux parties qui correspondent, en toute première approximation, aux cours d'algèbre enseignés respectivement au premier et au troisième semestre. L'étude de la partie II suppose que le lecteur a assimilé la théorie des espaces vectoriels abstraits et des opérateurs linéaires, enseignée au deuxième semestre dans le cours d'algèbre linéaire et de géométrie. D'ailleurs, les espaces vectoriels $\mathbb{R}^n$ des vecteurs lignes sont exposés au chapitre 2, certaines notions d'algèbre linéaire sont introduites dans le texte du livre chaque fois que le besoin s'en fait sentir, et un petit annexe donné à la fin du livre expose la théorie géométrique de la réduction des matrices à la forme de Jordan. Ainsi, le présent manuel peut être étudié indépendamment des autres sources.

Un rôle bien important est dévolu aux exercices donnés en fin de la plupart des paragraphes. Vu l'existence d'excellents ouvrages consacrés aux problèmes d'algèbre, il n'a pas paru raisonnable de mettre l'accent sur les calculs numériques, si bien que les exercices figurant dans ce manuel ont essentiellement un caractère théorique et servent au développement du sujet principal. Les exercices auxquels on fait référence dans le texte principal, sont munis d'indications détaillées pour leur résolution. Nous conseillons au lecteur de se reporter à ces indications aussi rarement que possible, et seulement si ses tentatives personnelles de trouver la solution correcte ne donnent pas de résultat.

Il est difficile de s'attendre que le nombre d'heures assez modeste réservé aux conférences soit suffisant pour recouvrir tout le contenu

du livre. Cela concerne surtout la partie II dont le matériel ne peut pas être considéré, d'après son caractère, comme traditionnel. Ce matériel donne matière à l'intuition, mais certains «mets délicats » (tels les théorèmes de Sylow, les invariants des groupes linéaires, les représentations du groupe des rotations ou les algèbres non associatives) sont destinés exprès aux amateurs comme base d'études supplémentaires.

Il paraît qu'après l'étude du chapitre 7 assez difficile, il convient de s'orienter soit vers les éléments de théorie des représentations (chapitre 8), soit vers la théorie générale des anneaux, modules et corps traitée en partie dans le chapitre 9 (une étude approfondie de la théorie des structures algébriques sortirait nettement du cadre du présent ouvrage). La première variante semble être préférable, et ceci non seulement par suite de son orientation géométrique et de sa proximité du cours d'algèbre linéaire et de géométrie enseigné au deuxième semestre, mais aussi parce que la connaissance des faits principaux de la théorie des représentations des groupes est très utile même pour les mathématiciens qui ne se spécialisent pas en algèbre. Il est extrêmement souhaitable que l'idée de représentations des groupes exprimée dans ce livre sur un matériel concret peu important soit corroborée dans un cours spécial plus solide. Comme sujets de ce cours, on pourrait indiquer, par exemple, la théorie de Galois, les groupes engendrés par les symétries orthogonales, y compris les groupes cristallographiques, les représentations des groupes compacts, etc. D'autre part, le chapitre 9 qui met l'accent sur la théorie des nombres, répond davantage aux programmes en vigueur. Quelle que soit la variante choisie, elle jettera les bases de l'étude ultérieure de l'algèbre *).

Maintenant, il convient de signaler une circonstance qui n'est pas trop évidente pour un étudiant débutant. Malgré son appellation, un cours d'algèbre supérieure est loin de refléter toute la variété de l'algèbre moderne. C'est pour cette raison que le présent livre est intitulé « Introduction à l'algèbre ». L'un des objectifs de cet ouvrage est de fournir des notions et des résultats nécessaires pour d'autres cours de mathématiques. On ne peut comprendre l'importance du langage mathématique qu'en essayant de s'en passer lors d'une étude individuelle des mathématiques.

Malgré son caractère élémentaire, un cours d'algèbre traditionnel présente certaines difficultés quant à son assimilation, à cause de la pensée abstraite dont il exige de faire la preuve. Ayant constamment en vue cette circonstance, l'auteur a cherché à souligner les liens qui existent entre l'algèbre et d'autres branches des mathématiques. Il est regrettable que les chapitres consacrés aux éléments

*) Au début de chaque partie du livre on trouvera une bibliographie sommaire.

de théorie des catégories et des systèmes partiellement ordonnés n'aient pas trouvé de place dans le présent ouvrage. Cependant, il n'a pas paru raisonnable d'assimiler un cours d'introduction à un conglomérat de notions abstraites qu'on introduit à profusion dans un but inconnu et qui tuent tout intérêt parce que leur étude est trop sommaire.

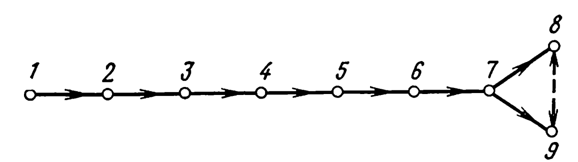
De nombreuses variantes concevables de cours obligatoire d'algèbre, limité et orienté par le programme standard, ont été pratiquement essayées à la faculté de Mécanique et de Mathématiques de l'Université de Moscou. On peut espérer que la présente réalisation, sous forme d'un livre, de l'une des dernières variantes de ce cours sera utile pour les étudiants et les enseignants des autres écoles supérieures, ainsi que pour ceux qui commencent à étudier individuellement le cours d'algèbre. Bien entendu, l'ordre et la plénitude de l'exposé du contenu de ce livre au cours des conférences dépendront fortement de la situation concrète et des traditions d'enseignement qui se sont établies.

L'auteur est très obligé aux professeurs de la chaire d'algèbre supérieure de l'Université de Moscou et tient à remercier ici tous ceux qui ont bien voulu lui donner les précieux conseils concernant l'exposé du cours. Toutes les suggestions constructives, ainsi que les avis sur des incorrections constatées seront accueillis avec reconnaissance.

Moscou, juin 1976. *A. KOSTRIKIN*

AVIS AU LECTEUR

Suivant le plan général de l'ouvrage exposé dans la préface, l'enchaînement logique des chapitres se présente ainsi:



(la flèche pointillée indique une faible dépendance). Il est clair qu'un lecteur expérimenté (disons, un professeur de mathématiques ou même un étudiant en deuxième année) n'aura pas de grandes difficultés à commencer la lecture de ce manuel pratiquement par n'importe quel endroit, certes, s'il est prêt à se reporter, au fur et à mesure des besoins, aux définitions données aux paragraphes et chapitres précédents. Les nouvelles notions ne sont pas toutes introduites dans des alinéas commençant par le mot « Définition ». La table des matières détaillée et l'index alphabétique faciliteront la recherche de l'endroit correspondant dans le livre.

Chaque chapitre est subdivisé en plusieurs paragraphes dont chacun comprend plusieurs numéros ayant leurs propres titres. Les théorèmes, propositions, lemmes et corollaires ont leur propre numération à l'intérieur de chaque paragraphe: théorème 1, théorème 2, . . .; lemme 1, lemme 2, . . . Avec une telle numération, les références aux assertions énoncées dans un autre paragraphe se font ainsi: théorème i du § j, ou même, théorème i du chapitre k, § j, mais cela ne cause pas d'incommodités.

La fin d'une démonstration (ou son absence) est marquée par le signe $\blacksquare$.

Pour abréger l'exposé on utilise des symboles logiques simples. Le symbole d'implication $\Rightarrow$ dans la notation $A \Rightarrow B$ signifie « A implique (ou entraîne) B », alors que « $A \Leftrightarrow B$ » se lit « A est équivalent à B ». Le quantificateur universel $\forall$ sert à représenter l'expression « Quel que soit » ou « Pour tout ». Les autres désignations et symboles seront expliqués par le contexte.

Pour éviter toute confusion dans l'emploi des lettres grecques, très usitées en mathématiques, nous rappelons ci-dessous cet alphabet *in extenso*:

ALPHABET GREC

Α α	a	alpha	Ν ν	n	nu
Β β	b	bêta	Ξ ξ	ks	xi
Γ γ	g	gamma	Ο ο	o	omicron
Δ δ	d	delta	Π π	p	pi
Ε ε	e	epsilon	Ρ ρ	r	rhô
Ζ ζ	dz	dzéta	Σ σ	s	sigma
Η η	e	êta	Τ τ	t	tau
Θ θ	t	aspiré : thêta	Υ υ	u	upsilon
Ι ι	i	iota	Φ φ	p	aspiré : phi
Κ κ	k	kappa	Χ χ	k	aspiré : khi
Λ λ	l	lambda	Ψ ψ	ps	psi
Μ μ	m	mu	Ω ω	o	oméga

PARTIE I

NOTIONS FONDAMENTALES D'ALGÈBRE

Cette partie peut être considérée comme une algèbre en miniature. Les notions fondamentales de groupe, d'anneau, de corps, inhabituelles pour un étudiant débutant, sont introduites, autant qu'il est possible, de façon non formelle et dans des doses minimales, bien que le nombre de notions dérivées soit assez grand. On ne doit pas chercher à les apprendre « par cœur »: elles deviendront familières après un travail individuel sur les problèmes et exercices. Pour rendre plus commode l'exposé, on dégage quelques systèmes algébriques, les plus utilisés (groupes $(\mathbb{Z}, +)$, S_n, A_n, GL (n), SL (n); anneau de polynômes; corps $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ et Z_p) et, sur leur base, on introduit et explique le langage algébrique. Conformément à la tradition et pour assurer une continuité entre l'enseignement secondaire et l'enseignement supérieur, on expose au début la technique des matrices et des déterminants, qui est utilisée pour la recherche et la discussion des solutions de systèmes d'équations linéaires. C'est dans cette voie qu'apparaissent de façon naturelle les structures algébriques fondamentales.

BIBLIOGRAPHIE

1. Kurosh A., *Cours d'algèbre supérieure*, Ed. Mir, Moscou, 1980.
2. Leng S., *Algebra*, Columbia University, New York, Addison-Wesley publishing company Reading mass, 1965.
3. Van der Waerden B. L., *Algebra* I, Achte Auflage der modernen Algebra, Springer-Verlag, Berlin, Heidelberg, New York, 1971. *Algebra* II, Fünfte Auflage, Springer-Verlag, Berlin, Heidelberg, New York, 1967.
4. Виноградов И. М., *Основы теории чисел*, « Наука », 1972 (Vinogradov I., *Eléments de théorie des nombres*, Moscou, 1972).
5. Проскуряков И. В., *Сборник задач по линейной алгебре*, « Наука », 1974 (Proskouriakov I., *Recueil de problèmes d'algèbre linéaire*, Moscou, 1974).

CHAPITRE PREMIER

LA GENÈSE DE L'ALGÈBRE

Par quoi commence l'algèbre? On peut dire avec une certaine approximation que l'algèbre a son origine dans l'art d'additionner, de multiplier et d'élever à une puissance les nombres entiers. Le remplacement des nombres par les lettres, quoique formel, mais loin d'être évident et univoque, permet d'opérer suivant les règles analogues dans les limites des systèmes algébriques considérablement plus généraux. Cela signifie que la tentative de donner une réponse complète à la question posée nous conduirait non seulement dans la profondeur des siècles mais aussi au mystère de la naissance de la pensée mathématique. La partie la plus difficile de la réponse serait liée à la description des structures fondamentales de l'algèbre moderne: des groupes, des anneaux, des corps, des modules, etc. Or, c'est justement à cette description qu'est consacré tout le livre, de sorte que le but du chapitre 1 semble pour l'instant impossible à atteindre.

Heureusement, sous l'enveloppe abstraite de la plupart des théories axiomatiques de l'algèbre, on trouve des problèmes tout à fait concrets de caractère théorique ou pratique dont la résolution donnait lieu à des généralisations heureuses et, parfois inévitables, de très grande portée. A son tour, une théorie bien développée incitait à poser de nouveaux problèmes et fournissait des moyens permettant leur résolution. Cette interaction complexe entre l'aspect théorique et l'aspect appliqué, inhérente à toutes les branches de mathématiques, se manifeste très nettement en algèbre et justifie, dans une certaine mesure, la méthode concentrique de l'exposition que nous avons adoptée dans le présent ouvrage.

Après de brèves remarques générales relatives à l'historique de l'algèbre, nous énoncerons quelques problèmes qui prédéterminent en quelque sorte le contenu des chapitres suivants. Un de ces problèmes servira de point de départ pour l'étude des systèmes d'équations linéaires, de la théorie des matrices et de la théorie des déterminants. Nous décrirons la méthode de Gauss et obtiendrons les premiers résultats relatifs à la résolution des systèmes linéaires.

Dans cette étape même, il est utile d'introduire le vocabulaire et les symboles que les mathématiques mettent constamment en œuvre; à cet effet, nous donnerons un aperçu sommaire de la théorie des ensembles et des applications.

Nous introduirons aussi les notions importantes de relation d'équivalence et de factorisation des applications. Puis, en expliquant le principe d'induction mathématique (de récurrence), nous établirons des relations combinatoires élémentaires. Enfin, les propriétés arithmétiques les plus simples du système des nombres entiers, que nous décrirons dans le dernier paragraphe, seront non seulement utilisées par la suite, mais serviront de modèle pour la construction d'une arithmétique analogue dans des systèmes algébriques plus complexes.

Le contenu de ce chapitre ne dépasse pas trop le programme de l'enseignement secondaire. Il faut seulement que le lecteur soit prêt à accepter un point de vue plus général. On peut commencer la lecture du livre à partir du § 3.

§ 1. Aperçu historique

Les mathématiques de nos jours portent à juste titre le nom de « mathématiques algébrisées » puisqu'il s'agit là d'une pénétration des idées et des méthodes algébriques dans toutes les branches des mathématiques tant pures qu'appliquées. Un tel état de choses, devenu évident vers le milieu du XXe siècle, n'était pas toujours le même. Comme tout autre domaine de l'activité humaine, les mathématiques sont sujettes à l'influence de la mode qui, guidée par des besoins de la science, dictait parfois un emploi exagéré des méthodes algébriques. C'est pourquoi, vu qu'une enveloppe algébrique qui masque le contenu, est un mal non moindre qu'un simple oubli de l'algèbre, on considère (à juste titre), comme un avantage d'un livre, le fait qu'il n'est pas surchargé de formalisme algébrique.

Sans porter tout à l'extrême, on peut dire que l'algèbre était toujours l'une des parties essentielles des mathématiques. Il faudrait dire la même chose de la géométrie, mais il vaut mieux citer l'expression imagée due à Sophie Germain (XIXe s.) qui disait que l'algèbre n'était rien d'autre que la géométrie écrite en symboles, et la géométrie, c'était simplement l'algèbre incarnée dans les figures. Depuis, la situation a changé, mais, comme l'affirme N. Bourbaki, il paraît qu'on a reconnu que la « nature » des êtres mathématiques est, au fond, une chose secondaire et il importe peu, par exemple, que nous représentions le résultat sous la forme d'un théorème de la géométrie « pure » ou à l'aide de la géométrie analytique sous la forme d'un théorème algébrique.

Conformément au principe « Ce sont les relations entre des êtres mathématiques qui importent et non pas les êtres eux-mêmes » l'algèbre peut être définie (d'une façon un peu tautologique et tout

à fait incompréhensible pour un non initié) comme la science ayant pour objet l'étude des opérations algébriques effectuées sur les éléments de divers ensembles. Les opérations algébriques elles-mêmes sont issues de l'arithmétique élémentaire. D'autre part, en partant des considérations algébriques, on donne les démonstrations les plus naturelles de nombreuses propositions de l'« arithmétique supérieure », c'est-à-dire de la théorie des nombres.

L'importance que présentent les structures algébriques, c'est-à-dire les ensembles munis d'une ou de plusieurs opérations algébriques, ne réside pas uniquement dans leurs applications à la théorie des nombres. L'étude de nombreux êtres mathématiques (espaces topologiques, équations différentielles, fonctions de plusieurs variables complexes et autres) se fait par construction de structures algébriques correspondantes qui, bien qu'elles ne soient pas adéquates à des objets à étudier, reflètent leurs traits essentiels. Quelque chose de pareil s'applique aussi aux objets du monde réel.

Une opinion bien déterminée à ce sujet a été exprimée il y a une cinquantaine d'années par Paul Dirac, l'un des créateurs de la mécanique quantique, qui disait que la physique moderne exigeait une mathématique de plus en plus abstraite et le développement de ses fondements. C'est ainsi que la géométrie non euclidienne et l'algèbre non commutative qui étaient considérées autrefois comme de pures imaginations, sont maintenant reconnues bien nécessaires à la description générale de la réalité physique.

Les moyens algébriques se montrent bien utiles pour l'étude des particules élémentaires en mécanique quantique, pour l'étude des propriétés des corps solides et des cristaux (c'est la théorie des représentations des groupes qui joue ici un rôle particulièrement important), pour l'analyse des problèmes d'économie types, pour la conception des calculateurs électroniques, etc.

A son tour, l'algèbre met à profit les méthodes et les idées développées dans d'autres sciences, y compris les disciplines mathématiques. C'est ainsi, par exemple, que les méthodes homologiques en algèbre sont nées de la topologie et de la théorie algébrique des nombres. Ce n'est pas donc étonnant que l'aspect de l'algèbre et la conception de cette branche des mathématiques changeaient d'une époque à l'autre. Nous n'avons pas la possibilité de suivre pas à pas ces changements, et ceci non seulement parce que la place nous manque, mais surtout parce que la description de l'histoire d'une matière doit être concrète ; or, on ne peut faire face à cette exigence qu'après une étude approfondie de cette matière. Nous nous contenterons donc d'énumérer schématiquement les noms et les périodes.

Civilisations antiques de Babylone et d'Egypte. Civilisation grecque. « Arithmétiques » de Diophante d'Alexandrie (III[e] s. n.è.)	Opérations arithmétiques sur les nombres entiers et les nombres rationnels positifs. Formules algébriques dans les calculs géométriques et astronomi-

	ques. Enoncé de problèmes de construction (duplication du cube, trisection de l'angle) qui ont occupé sensiblement plus tard les esprits des algébristes.
Civilisation orientale de Moyen Age. Ouvrage de l'originaire de Khiva Muhammed ibn Musà al-Khawàrizmi (vers 825) « Hisab-al-jabr wal-muquàbala ».	Equations algébriques du premier et du second degré. Apparition du terme « algèbre ».
Renaissance.	Résolution des équations algébriques générales du troisième et du quatrième degré.
Leonardo Fibonacci (dit « Pisano ») (vers 1170-1250) Spicione dal Ferro (1465-1526) N. Tartaglia (1500-1557) H. Cardan (1501-1576) L. Ferrari (1522-1565) F. Viète (1540-1603) R. Bombelli (1530-1572)	Création du symbolisme algébrique moderne.
XVIIe et XVIIIe s. R. Descartes (1596-1650) P. Fermat (1601-1665) I. Newton (1643-1727) G. Leibniz (1646-1716) L. Euler (1707-1783) J. d'Alembert (1717-1783) J. L. Lagrange (1736-1813) G. Cramer (1704-1752) P. Laplace (1749-1827) A. Vandermonde (1735-1796)	Création de la géométrie analytique constituant un pont solide entre la géométrie et l'algèbre. Renouveau d'activité dans la théorie des nombres. Développement de l'algèbre des polynômes. Recherches intenses des formules générales pour la résolution des équations algébriques. Premières approches de la démonstration de l'existence de racines d'une équation à coefficients numériques. Débuts de la théorie des déterminants.
XIXe s.-début du XXe s. C. F. Gauss (1777-1855) P. G. L. Dirichlet (1805-1859) E. Kummer (1810-1893) L. Kronecker (1823-1891) R. Dedekind (1831-1916) E. I. Zolotarev (1847-1878) G. F. Voronoï (1868-1908) A. A. Markov (1856-1922)	Démonstration du théorème fondamental d'existence des solutions des équations à coefficients numériques. Développement intense de la théorie des nombres algébriques.
P. L. Tchébychev (1821-1894) Ch. Hermite (1822-1901) N. I. Lobatchevski (1792-1856) A. Hurwitz (1859-1919)	Recherche des méthodes permettant d'obtenir des solutions approchées des équations algébriques. Conditions imposées aux coefficients pour assurer une disposition donnée des racines.
A. Ruffini (1765-1822) N. H. Abel (1802-1829) C. Jacobi (1804-1851) E. Galois (1811-1832) B. Riemann (1826-1866)	Démonstration de l'irrésolubilité par radicaux des équations générales du degré $n \geqslant 5$. Développement de la théorie des fonctions algébriques. Création de la théorie de Galois.

A. Cauchy (1789-1857) C. Jordan (1838-1922) L. Sylow (1832-1918)	Début de la théorie des groupes finis, basée essentiellement sur les groupes de permutations.
H. Grassmann (1809-1877) J. Sylvester (1814-1897) A. Cayley (1821-1895) W. Hamilton (1805-1865) G. Boole (1815-1864) S. Lie (1842-1899) G. Frobenius (1849-1918) J.-A. Serret (1819-1895) M. Nöther (1844-1922) D. A. Gravet (1863-1939) H. Poincaré (1854-1912) F. Klein (1849-1925) W. Burnside (1852-1927) I. Schur (1885-1941) H. Weyl (1885-1955) F. Enriques (1871-1946) J. Neumann (1903-1957) D. Hilbert (1862-1943) E. Cartan (1869-1951) K. Hensel (1861-1941) E. Steinitz (1871-1928) E. Noether (1882-1935) E. Artin (1898-1962) N. Bourbaki « Eléments de mathématique »	Développement intense des méthodes de l'algèbre linéaire. Apparition, après la découverte des quaternions, de la théorie des systèmes hypercomplexes (qu'on appelle actuellement algèbres). En particulier, grâce au développement de la théorie des groupes continus (groupe de Lie) on jette les bases de la théorie des algèbres de Lie. La géométrie algébrique et la théorie des invariants deviennent des chapitres importants des mathématiques. Au XIXe siècle, les mathématiques n'ont pas encore atteint un haut degré de différenciation, de sorte que de nombreux savants éminents travaillent dans plusieurs branches. La première moitié du XXe siècle a été marquée par une refonte complète de tout l'édifice mathématique. L'algèbre a renoncé au privilège d'être science des équations algébriques et a résolument emprunté pour son développement une voie axiomatique et sensiblement plus abstraite.

Le langage de la théorie des anneaux, des modules, des catégories et des homologies a pénétré dans l'usage des mathématiciens. Plusieurs théories séparées sont réunies dans le cadre d'un schéma général de l'algèbre universelle. Une théorie des modèles est née à la charnière de l'algèbre et de la logique mathématique. De vieilles théories ont connu un renouveau et un élargissement du domaine de leurs applications. On peut citer à titre d'exemple la géométrie algébrique moderne, la topologie algébrique, la K-théorie algébrique et la théorie des groupes algébriques. La théorie des groupes finis a connu quelques essors éclatants.

Toute l'algèbre est actuellement en état de développement dynamique. De grands mérites reviennent aux mathématiciens soviétiques. Le haut niveau de recherches algébriques en U.R.S.S. est dû pour une large part aux savants tels que N. G. Tchébotarev (1894-1947), O. J. Schmidt (1891-1956), A. I. Maltsev (1909-1967), A. G. Kurosh (1908-1971), P. S. Novikov (1901-1975).

§ 2. Quelques problèmes types

Les quatre problèmes que nous allons énoncer ci-dessous se placent à des niveaux différents. Les trois premiers, qui ne sont pas non plus équivalents l'un à l'autre, sont destinés uniquement à motiver l'étude de divers types de corps, des espaces vectoriels, des groupes

et de leurs représentations, c'est-à-dire l'étude des théories algébriques dont il va s'agir par la suite. De nombreuses monographies spéciales sont consacrées à la « résolution » de ces problèmes. Quant au quatrième problème qui anticipe sur l'étude des systèmes linéaires, il est utile d'essayer de le résoudre tout de suite, sans consulter le paragraphe suivant où est développé le raisonnement nécessaire.

1. Problème de résolubilité des équations par radicaux. — On connaît la formule obtenue en algèbre élémentaire

$$x_{1,2}=\frac{-b\pm\sqrt{b^2-4ac}}{2a} \qquad (1)$$

qui donne les solutions x_1, x_2 de l'équation du second degré $ax^2 + bx + c = 0$.

L'équation du troisième degré $x^3 + ax^2 + bx + c = 0$ est ramenée, moyennant la substitution $x \mapsto x - \frac{1}{3}a$, à la forme $x^3 + px + q = 0$. Cette équation réduite possède toujours trois racines x_1, x_2, x_3. Si l'on pose

$$D=-4p^3-27q^2, \qquad \varepsilon=\frac{-1+\sqrt{-3}}{2},$$
$$u=\sqrt[3]{-\frac{27}{2}q+\frac{3}{2}\sqrt{-3D}}, \qquad v=\sqrt[3]{-\frac{27}{2}q-\frac{3}{2}\sqrt{-3D}} \qquad (2)$$

(les racines cubiques sont choisies de façon que $uv = -3p$), on peut montrer que

$$x_1=\frac{1}{3}(u+v), \quad x_2=\frac{1}{3}(\varepsilon^2u+\varepsilon v), \quad x_3=\frac{1}{3}(\varepsilon u+\varepsilon^2 v). \qquad (3)$$

Les formules (2) et (3) appelées formules de Cardan (1545) et s'associant aussi aux noms de certains autres mathématiciens italiens de la Renaissance (S. Ferro, N. Tartaglia) sont valables, de même que la formule (1), quels que soient les coefficients littéraux a, b, c, p, q auxquels on peut donner par exemple des valeurs rationnelles arbitraires. Des formules analogues ont été obtenues pour les racines de l'équation du quatrième degré, mais c'est en vain qu'on a essayé, pendant près de trois cents ans, de « résoudre par radicaux » l'équation générale du cinquième degré. Ce n'est qu'en 1813 que A. Ruffini (en première approximation) et en 1827 N. Abel (indépendamment de Ruffini et de façon tout à fait rigoureuse) ont démontré le théorème énonçant qu'il était impossible de résoudre par radicaux l'équation générale $x^n + a_1x^{n-1} + \ldots + a_n = 0$, pour $n > 4$. Une découverte fondamentale dans ce domaine a été faite en 1831 (elle n'a été publiée qu'en 1846) par Evariste Galois, à l'âge de vingt ans, quand il a donné les conditions nécessaires et suffisantes pour que, non seulement une équation générale de degré n, mais n'importe quelle

équation (par exemple à coefficients rationnels) soit résoluble par radicaux.

A chaque polynôme (équation) de degré n, il a fait correspondre un corps de décomposition et une famille finie (de puissance non supérieure à $n!$) d'automorphismes de ce corps qu'on appelle maintenant groupe de Galois du corps (ou du polynôme de base). Bien que nous n'ayons pas la possibilité d'étudier la théorie de Galois avec plus de détails, nous dégagerons au chapitre 7, à partir des propriétés purement internes, une classe spéciale de groupes dits résolubles. Il se trouve qu'une équation de degré n à coefficients rationnels est résoluble par radicaux si, et seulement si, est résoluble le groupe de Galois qui correspond à cette équation. Soit donnée, par exemple, l'équation du cinquième degré $x^5 - ax - 1 = 0$, où a est un entier. A cette équation correspond un groupe de Galois G_a qui dépend de a suivant une certaine loi complexe. G_0 étant un groupe cyclique d'ordre 4 (tous les groupes cycliques sont résolubles par définition), l'équation $x^5 - 1 = 0$ est donc résoluble par radicaux. Par contre, G_1 est de même structure que le groupe symétrique S_5 d'ordre 120 et, comme il sera montré au chapitre 7, ce dernier est irrésoluble. Par suite, l'équation $x^5 - x - 1 = 0$ est, elle aussi, irrésoluble par radicaux.

Notons, avant de clore la discussion de ce problème, que la possibilité d'exprimer la racine d'une équation algébrique par radicaux, sous forme explicite, est sans grande importance pratique ; ce sont les diverses méthodes approchées de calcul des racines qui offrent un plus grand intérêt pratique. Mais cette circonstance ne compromet nullement l'élégance des résultats obtenus par Galois dont les idées ont exercé une forte influence sur le développement ultérieur des mathématiques. Il suffit de dire que Galois a jeté les bases de la théorie des groupes. La correspondance biunivoque, établie par E. Galois, entre les sous-corps du corps de décomposition et les sous-groupes du groupe de Galois s'est enrichie au XXe siècle de nouvelles structures abstraites et est devenue un instrument indispensable pour l'étude des êtres mathématiques.

2. Problème sur les états d'une molécule polyatomique.— Toute molécule peut être considérée comme un système de particules qui sont des noyaux atomiques (entourés d'électrons). Si, à l'instant initial, la configuration du système est voisine de celle en équilibre, les particules constitutives du système resteront toujours, dans des conditions déterminées, près de la position d'équilibre et ne seront pas animées de grandes vitesses. Les mouvements de ce type sont appelés oscillations autour de la configuration d'équilibre, et le système lui-même est dit stable. On sait que toute petite oscillation d'une molécule autour de sa position d'équilibre stable est une superposition des oscillations dites normales. Dans de nombreux

cas, on arrive à déterminer l'énergie potentielle de la molécule et ses fréquences normales en tenant compte de la symétrie interne de la molécule. Cette dernière se décrit par le groupe ponctuel de la molécule. Les diverses réalisations de ce groupe fini (ses représentations irréductibles) et les fonctions sur le groupe liées à ces réalisations (caractères des représentations) déterminent les paramètres d'oscillations de la molécule.

C'est ainsi par exemple qu'à une molécule d'eau H_2O (fig. 1) correspond un groupe de Klein à quatre éléments (produit direct de deux groupes cycliques du deuxième ordre), alors qu'à une molécule de phosphore P_4 (fig. 2), ayant l'aspect d'un tétraèdre régulier

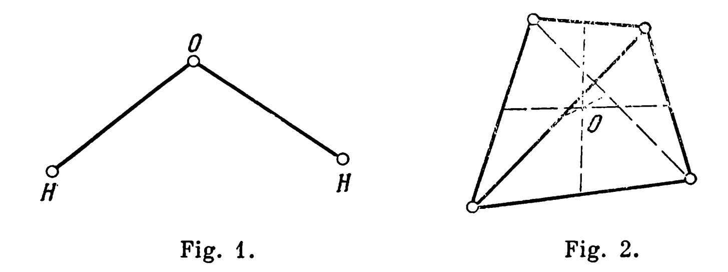


Fig. 1. Fig. 2.

aux sommets duquel se situent les atomes de phosphore, correspond un groupe symétrique S_4 d'ordre 24. Les représentations irréductibles de ces groupes seront étudiées au chapitre 8. De nos jours, il est difficile de concevoir le développement de la théorie des structures moléculaires sans aide de la théorie des groupes.

Les applications beaucoup plus antérieures de la théorie des groupes se rapportent à la cristallographie. Dès 1891, le grand cristallographe russe E. S. Fedorov et ensuite le savant allemand A. Schönflies ont trouvé 230 groupes cristallographiques d'espace qui décrivent toutes les symétries des cristaux existant dans la nature. Depuis, la théorie des groupes est constamment utilisée pour étudier l'influence de la symétrie sur les propriétés physiques des cristaux.

3. Problème de codage d'un message.— Lors de l'élaboration des systèmes de télécommunications automatiques, terrestres ou cosmiques, on prend généralement en qualité de message élémentaire une suite ordonnée, c'est-à-dire une ligne (ou un mot) $a = (a_1, a_2, \ldots, a_n)$ de longueur n, où $a_i = 0$ ou 1. Vu que les opérations ordinaires d'addition et de multiplication modulo 2 sont bien faciles à réaliser sur un calculateur électronique et que les symboles 0, 1 peuvent être commodément transmis à l'aide de signaux électriques (1 et 0 diffèrent par la phase des signaux partagés dans le temps, ou bien par leur présence ou leur absence), ce n'est pas étonnant que le corps GF (2) (voir chap. 4, § 4) est un outil

indispensable entre les mains d'un spécialiste de traitement de l'information. Parfois, il est commode d'utiliser comme a_i les éléments appartenant à d'autres corps finis.

Pour éliminer l'influence des parasites (décharges atmosphériques bruits cosmiques, etc.) capables de transformer 0 en 1 et vice versa, on est amené à prendre a d'une longueur suffisamment grande et à utiliser un système de *codage* spécial, c'est-à-dire à choisir dans un ensemble S de lignes le sous-ensemble (un *code*) S_0 des lignes à transmettre (des mots de code), de manière qu'il soit possible de rétablir a d'après le mot déformé reçu a', à condition que le nombre d'erreurs intervenues ne soit pas trop grand. C'est ainsi qu'on élabore des *codes à correction d'erreurs.* La théorie algébrique du codage s'est fortement développée ces dernières années et a proposé de nombreuses méthodes ingénieuses de codage. Elle a surtout affaire à des codes linéaires spéciaux pour lesquels le choix de S_0 est lié à la construction de matrices rectangulaires spéciales et à la résolution de systèmes d'équations linéaires dont les coefficients appartiennent à un corps fini donné. Un exemple simple d'un tel code sera donné au chapitre 5.

4. Problème de la plaquette chauffée.— Une plaquette plane de forme rectangulaire présentant trois orifices (fig. 3) est utilisée comme soupape dans un dispositif fantastique destiné à obtenir des

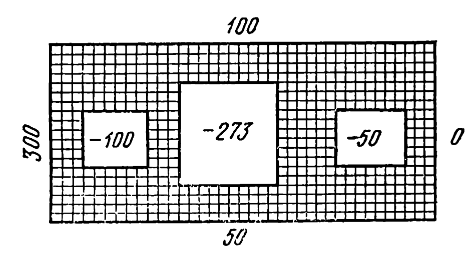


Fig. 3.

basses températures. Un réseau carré (un quadrillage) est posé sur la soupape. Les nœuds du réseau situés sur les quatre contours s'appellent nœuds frontières, tous les autres nœuds étant intérieurs. Une mesure directe montre que lors de tout échauffement ou de tout refroidissement, la température en chaque nœud intérieur est la moyenne arithmétique des températures des quatre nœuds les plus proches, qu'ils soient frontières ou intérieurs. On s'attend qu'étant mises en contact avec les différentes portions des contours, les pièces constitutives du dispositif communiqueront aux points frontières les températures indiquées sur la fig. 3. Est-ce possible ? Si oui, la répartition des températures aux points intérieurs sera-t-elle univoque ?

§ 3. Systèmes d'équations linéaires. Premiers pas

Les équations linéaires $ax = b$ et les systèmes de la forme

$$\begin{aligned} ax + by &= e, \\ cx + dy &= f \end{aligned} \tag{1}$$

à coefficients réels a, b, c, d, e, f sont « résolus » à l'école secondaire. Notre but est d'apprendre à opérer sur les *systèmes d'équations algébriques linéaires* (en abrégé, *systèmes linéaires*) de la forme la plus générale

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m, \end{aligned} \tag{2}$$

où m et n sont des entiers positifs arbitraires. Les complications qui naissent lors du passage de (1) à (2) et qui ont l'air d'être de nature purement quantitative, présentent en réalité une importance de principe. Les systèmes de la forme (2) se rencontrent littéralement dans toutes les branches des mathématiques, et les méthodes dites linéaires, dont le produit final représente souvent les solutions des systèmes linéaires, constituent la partie la plus développée des mathématiques. Il suffit de mentionner que la théorie des systèmes de la forme (2) a servi de base pour l'élaboration, à la fin du XIX[e] siècle, de la théorie des équations intégrales qui joue un rôle extrêmement important en mécanique et en physique. La résolution de nombreux problèmes pratiques sur calculateur électronique se ramène, elle aussi, à la résolution des systèmes (2).

1. Terminologie.— On remarquera la désignation bien commode et fort économique des coefficients dans le système (2) : le *coefficient* a_{ij} (se lit a-i-j ; par exemple a_{12} se lit a-un-deux et non pas a-douze) se trouve dans la i-ième équation devant la j-ième *inconnue* x_j. Le nombre b_i s'appelle *terme constant* de la i-ième équation. On dit que le système (2) est homogène si $b_i = 0$ pour $i = 1, 2, \ldots, m$. Quels que soient les b_i, le système linéaire

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= 0, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= 0 \end{aligned} \tag{2°}$$

est appelé *système homogène associé* au système (2) ou encore *forme réduite* du système (2).

Les coefficients des inconnues forment un tableau rectangulaire

$$\left\|\begin{matrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{matrix}\right\| \tag{3}$$

appelé *matrice* $m \times n$ (on dit aussi *matrice à m lignes et n colonnes* ou *matrice de type* (m, n), ou bien *matrice carrée d'ordre n* si $m=n$) et désigné par le symbole (a_{ij}) ou tout simplement par une lettre A. Il est naturel de parler de la i-ième *ligne* $(a_{i1}, a_{i2}, \dots, a_{in})$ de la matrice (3) et de sa j-ième *colonne*

$$\left\|\begin{matrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{matrix}\right\|$$

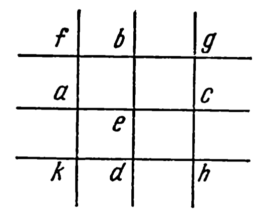
qui sera désignée par la suite, en vue d'économiser sur la place, par une ligne placée entre crochets : $[a_{1j}, a_{2j}, \dots, a_{mj}]$. Dans le cas d'une matrice carrée on parle encore de la *diagonale principale* formée par les éléments $a_{11}, a_{22}, \dots, a_{nn}$. Une matrice (a_{ij}) qui a tous ses éléments nuls, sauf ceux de la diagonale principale, est désignée parfois par le symbole diag $(a_{11}, a_{22}, \dots, a_{nn})$ et appelée *matrice diagonale*; lorsque $a_{11} = a_{22} = \dots = a_{nn} = a$, on la note $\mathrm{diag}_n(a)$ (*matrice scalaire*). Pour désigner la matrice $\mathrm{diag}_n(1)$ appelée *matrice unité*, on utilise généralement le symbole E_n ou une lettre E si les dimensions de la matrice sont fixées.

En plus de la matrice (3) on considère encore une matrice $(a_{ij} \mid b_i)$ du système (2), dite *complète*, qu'on obtient à partir de (3) par adjonction de la colonne $[b_1, b_2, \dots, b_m]$ de termes constants ; pour plus de clarté, cette colonne est séparée des autres par une barre verticale.

Si chacune des équations du système (2) se transforme en une identité après le remplacement des inconnues x_i par des nombres x_i°, la collection de n nombres $x_1^\circ, x_2^\circ, \dots, x_n^\circ$ s'appelle *solution* du système (2), alors que x_i° s'appelle i-ième *composante de la solution*. On dit encore que la collection $x_1^\circ, x_2^\circ, \dots, x_n^\circ$ satisfait à toutes les équations du système (2). Un système d'équations qui n'a aucune solution est dit *incompatible*. Si le système possède des solutions, on dit qu'il est *compatible*. Le système est déterminé si la solution est unique, et indéterminé quand il possède plusieurs solutions. Proposons-nous maintenant d'établir si un système donné d'équations linéaires est compatible et, s'il l'est, trouvons toutes ses solutions.

Revenons au problème du n° 4, § 2. Numérotons arbitrairement de 1 à 416 (tel est leur nombre sur la fig. 3) tous les points intérieurs de la plaquette, ajoutons à ces points 204 points frontières et, suivant la règle donnée pour le calcul

de la température t_i au point intérieur de numéro i, établissons 416 relations de la forme



$$t_e = \frac{t_a + t_b + t_c + t_d}{4}.$$

Supposons par exemple $a, b, c \leqslant 416$ et $d > 416$. Dans ce cas, la relation peut être mise sous forme d'une équation linéaire

$$-t_a - t_b - t_c + 4t_e = t_d$$

dont le second membre $t_d = -273, -100, -50, 0, 50, 100$ ou 300 (d'autres variantes sont également possibles). Prises ensemble, ces équations forment un système linéaire carré de la forme (2) avec $n = m = 416$. Les coefficients des inconnues t_i sont égaux à 0 (ils sont en majorité), à -1 ou à 4. Ce système est-il compatible et déterminé? Ainsi, nous avons obtenu un autre énoncé, mathématiquement correct, du problème de nature *qualitative*. Le problème d'existence et d'unicité de la solution est assez typique pour de nombreuses branches des mathématiques liées à l'étude des phénomènes physiques.

2. Equivalence des systèmes linéaires.— Soit donné encore un système linéaire « de même dimension » que (2):

$$\begin{aligned} a'_{11}x_1 + a'_{12}x_2 + \ldots + a'_{1n}x_n &= b'_1, \\ \cdots\cdots\cdots\cdots\cdots &\cdots \\ a'_{m1}x_1 + a'_{m2}x_2 + \ldots + a'_{mn}x_n &= b'_m. \end{aligned} \tag{2'}$$

Nous dirons que le système (2′) est obtenu à partir de (2) à l'aide d'une *transformation élémentaire de type* (I) si, dans le système (2′), toutes les équations sauf les i-ième et k-ième sont restées inchangées, alors que les i-ième et k-ième équations sont interverties entre elles. Si dans (2′) toutes les équations sauf la i-ième sont les mêmes que dans (2) et la i-ième équation a la forme

$$(a_{i1} + ca_{k1})\, x_1 + \ldots + (a_{in} + ca_{kn})\, x_n = b_i + cb_k, \tag{*}$$

où c est un nombre quelconque (c'est-à-dire $a'_{ij} = a_{ij} + ca_{kj}$, $b'_i = b_i + cb_k$), nous disons que le système (2) a subi une *transformation élémentaire de type* (II).

On dit que les systèmes (2) et (2′) sont *équivalents* s'ils sont soit incompatibles, soit compatibles et ont les mêmes solutions.

En convenant de désigner l'équivalence des systèmes (a) et (b) par le symbole $(a) \sim (b)$, nous remarquons que $(a) \sim (a)$; $(a) \sim (b)$ implique $(b) \sim (a)$; $(a) \sim (b)$ et $(b) \sim (c)$ entraînent $(a) \sim (c)$.

Une condition suffisante, pour que deux systèmes soient équivalents, est exprimée par le théorème suivant:

THÉORÈME 1. — *Deux systèmes linéaires sont équivalents si l'un d'eux est obtenu de l'autre par application d'une suite finie de transformations élémentaires.*

Il suffit de démontrer l'équivalence du système (2) et du système (2') obtenu du précédent par application d'une seule transformation élémentaire. Remarquons que le système (2) est obtenu de (2') aussi par application d'une seule transformation élémentaire, car ces transformations sont inversibles. En d'autres termes, lorsqu'il s'agit de la transformation de type (I), nous retrouvons le système initial après avoir permuté les équations de numéros i et k ; de façon analogue, dans le cas de la transformation de type (II), nous obtiendrons la i-ième équation du système (2) en ajoutant à la i-ième équation de (2') la k-ième équation multipliée par $(-c)$.

Démontrons maintenant que toute solution $(x_1^\circ, \ldots, x_n^\circ)$ du système (2) est aussi solution du système (2'). Si l'on a fait une transformation élémentaire de type (I), les équations, elles, ne sont pas changées en général (seul l'ordre de leur écriture est modifié). C'est pourquoi les nombres $x_1^\circ, x_2^\circ, \ldots, x_n^\circ$ qui satisfaisaient à ces équations avant la transformation, les vérifient aussi après la transformation. Dans le cas de la transformation élémentaire de type (II), les équations, sauf la i-ième, ne sont pas changées, si bien que la solution $(x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ leur satisfait comme précédemment. Quant à la i-ième équation, elle a pris la forme (*). Puisque notre solution satisfait aux i-ième et k-ième équations du système (2), on a

$$a_{i1}x_1^\circ + \ldots + a_{in}x_n^\circ = b_i, \quad a_{k1}x_1^\circ + \ldots + a_{kn}x_n^\circ = b_k.$$

Multiplions les deux membres de la dernière identité par c, ajoutons le résultat à la première et groupons les termes. Nous obtenons une identité de la forme (*) avec $x_i = x_i^\circ$.

Vu l'inversibilité des transformations élémentaires indiquée plus haut, le raisonnement développé montre également que, réciproquement, toute solution du système (2') sera aussi solution du système (2).

Il reste à remarquer que l'incompatibilité d'un système implique celle de l'autre (raisonnement par l'absurde). ■

3. Réduction à la forme quasi triangulaire.— Par l'application répétée des transformations élémentaires on peut passer d'un système d'équations donné à un système de forme plus simple.

Remarquons tout d'abord que parmi les coefficients a_{i1} l'un au moins est différent de zéro. Dans le cas contraire, il serait insensé de mentionner l'inconnue x_1. Si $a_{11} = 0$, permutons la première équation avec la j-ième, telle que $a_{j1} \neq 0$ (transformation de type (I)). Le coefficient de la première inconnue dans la première équation diffère maintenant de zéro. Désignons-le par a'_{11}. Retranchons de la i-ième équation $(i = 2, 3, \ldots, m)$ du nouveau système la première

équation dont les deux membres sont multipliés par un coefficient c_i, tel qu'après la soustraction le coefficient de x_1 s'annule ($m-1$ transformations élémentaires de type (II)). Il est évident qu'à cet effet il faut poser $c_i = a_{i1}/a'_{11}$. Il en résultera un système où x_1 n'intervient que dans la première équation. Il peut se trouver alors que la deuxième inconnue ne figure, elle non plus, dans aucune équation de numéro $i > 1$. Soit x_k l'inconnue de plus petit indice qui apparaît dans une équation quelconque, autre que la première. Nous obtenons ainsi le système

$$\begin{aligned} a'_{11}x_1 + \ldots\ldots\ldots\ldots + a'_{1n}x_n &= b', \\ a'_{2k}x_k + \ldots + a'_{2n}x_n &= b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a'_{mk}x_k + \ldots + a'_{mn}x_n &= b'_m, \quad k > 1, \quad a'_{11} \neq 0. \end{aligned}$$

Appliquons maintenant à toutes ces équations, sans faire attention à la première, les mêmes raisonnements que précédemment. Après une suite de transformations élémentaires, le système initial prend la forme

$$\begin{aligned} a''_{11}x_1 + \ldots\ldots\ldots\ldots\ldots\ldots + a''_{1n}x_n &= b''_1, \\ a''_{2k}x_k + \ldots\ldots\ldots + a''_{2n}x_n &= b''_2, \\ a''_{3l}x_l + \ldots + a''_{3n}x_n &= b''_3, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a''_{ml}x_l + \ldots + a''_{mn}x_n &= b''_m, \\ l > k > 1 \quad a''_{11} \neq 0, \quad a''_{2k} &\neq 0. \end{aligned}$$

Il va de soi qu'ici $a''_{1j} = a'_{1j}$, $b''_1 = b'_1$, car la première équation n'a subi aucune transformation.

Réitérons ce procédé tant qu'il sera possible de le faire. Il est clair qu'on devra s'arrêter à l'instant où deviendront nuls non seulement les coefficients de l'inconnue suivante (disons de la s-ième), mais aussi les coefficients de toutes les inconnues d'indice t, $s < t \leqslant n$. Finalement le système (2) prend la forme

$$\begin{aligned} \bar{a}_{11}x_1 + \ldots\ldots\ldots\ldots\ldots\ldots + \bar{a}_{1n}x_n &= \bar{b}_1, \\ \bar{a}_{2k}x_k + \ldots\ldots\ldots\ldots + \bar{a}_{2n}x_n &= \bar{b}_2, \\ \bar{a}_{3l}x_l + \ldots\ldots\ldots + \bar{a}_{3n}x_n &= \bar{b}_3, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ \bar{a}_{rs}x_s + \ldots + \bar{a}_{rn}x_n &= \bar{b}_r, \qquad (4) \\ 0 &= \bar{b}_{r+1}, \\ \ldots \\ 0 &= \bar{b}_m, \end{aligned}$$

avec $\bar{a}_{11}\bar{a}_{2k}\bar{a}_{3l} \ldots \bar{a}_{rs} \neq 0$, $1 < k < l < \ldots < s$. Il peut se trouver que $r = m$, de sorte que les équations de la forme $0 = \bar{b}_t$ ne seront pas présentées dans le système (4). On dit que le système d'équations (4) est présenté sous une forme *en échelons.*

Cette dénomination n'est pas universellement reconnue; on pourrait parler ici de la forme *trapézoïdale* ou *quasi triangulaire*, etc., mais cela n'est pas tellement important.

THÉORÈME 2. — *Tout système d'équations linéaires est équivalent à un système de forme quasi triangulaire.*

La démonstration découle immédiatement des raisonnements qui précèdent. ■

Parfois, il est plus commode d'appliquer les transformations élémentaires non pas au système lui-même, mais à sa matrice $(a_{ij} \mid b_i)$. Le théorème suivant se démontre exactement comme le théorème 2.

THÉORÈME 2'. — *Toute matrice peut être réduite à l'aide de transformations élémentaires à la forme quasi triangulaire.*

4. Discussion d'un système d'équations linéaires.— Du fait des théorèmes 1 et 2, il suffit d'étudier les questions de compatibilité et de détermination pour les systèmes de la forme quasi triangulaire (4).

Commençons par la compatibilité. Il est évident que si le système (4) contient une équation de la forme $0 = \bar{b}_t$, avec $\bar{b}_t \neq 0$, ce système est incompatible, parce que l'égalité $0 = \bar{b}_t$ ne peut pas être satisfaite quelles que soient les valeurs données aux inconnues. Démontrons que si le système (4) ne comporte pas de telles équations, il est compatible.

Soit $\bar{b}_t = 0$ pour $t > r$. Les inconnues $x_1, x_k, x_l, \ldots, x_s$ par lesquelles commencent respectivement les première, deuxième, . . ., r-ième équations, seront appelées inconnues *principales*, alors que les autres inconnues, si elles existent, seront dites *non principales* ou *libres.* Par définition, le nombre total d'inconnues principales est r.

Donnons aux inconnues non principales des valeurs arbitraires et introduisons-les dans les équations du système (4). On obtient alors pour x_s une seule équation (la r-ième) de la forme $ax_s = b$, avec $a = \bar{a}_{rs} \neq 0$, dont la solution est unique. On porte la valeur trouvée $x_s = x_s^\circ$ dans les $r - 1$ premières équations et on procède ainsi de proche en proche vers le haut du système (4). Il vient par suite que les valeurs des inconnues principales sont définies univoquement quelles que soient les valeurs données aux inconnues non principales. Nous avons ainsi démontré le théorème suivant:

THÉORÈME 3. — *Pour qu'un système d'équations linéaires soit compatible, il faut et il suffit qu'après la réduction à la forme quasi triangu-*

laire il ne contienne pas d'équations de la forme $0 = \bar{b}_t$, *avec* $\bar{b}_t \neq 0$. *Si cette condition est satisfaite, on peut donner aux inconnues non principales des valeurs arbitraires; les inconnues principales se définissent de façon univoque en fonction des valeurs données aux inconnues non principales.* ▣

Supposant satisfaite la condition de compatibilité que nous avons in troduite, proposons-nous maintenant d'étudier les conditions pour qu'un système soit déterminé. Si un système de la forme (4) contient des inconnues non principales, il est *a priori* indéterminé: nous pouvons donner aux inconnues non principales des valeurs quelconques et exprimer en fonction de ces dernières (suivant le théorème 3) les inconnues principales. Dans le cas où le système ne contient aucune inconnue non principale, et toutes les inconnues sont donc principales, le système est déterminé, car ses inconnues se définissent de façon univoque (théorème 3). Il reste à remarquer que l'absence d'inconnues non principales est équivalente à la condition $r = n$. Nous avons ainsi démontré l'assertion suivante:

THÉORÈME 4. — *Un système linéaire compatible* (2) *est déterminé, si, et seulement si, sa forme quasi triangulaire* (4) *vérifie l'égalité* $r=n$. ▣

Pour $m = n$, un système linéaire réduit à sa forme quasi triangulaire peut être mis aussi sous la forme (dite *triangulaire*)

$$\begin{aligned} \bar{a}_{11}x_1 + \bar{a}_{12}x_2 + \ldots + \bar{a}_{1n}x_n &= \bar{b}_1, \\ \bar{a}_{22}x_2 + \ldots + \bar{a}_{2n}x_n &= \bar{b}_2, \\ \ldots\ldots\ldots\ldots\ldots & \\ \bar{a}_{nn}x_n &= \bar{b}_n \end{aligned} \qquad (5)$$

si on ne se soucie pas de satisfaire pour tous les i la condition $\bar{a}_{ii} \neq 0$. En effet, (5) signifie que la k-ième équation du système ne contient pas des inconnues x_i, avec $i < k$; or, pour des systèmes ayant la forme quasi triangulaire cette condition est *a priori* satisfaite.

Notons encore, afin d'utiliser par la suite, qu'une matrice (a_{ij}) à éléments $a_{ij} = 0$, $i > j$, s'appelle matrice *triangulaire supérieure.* On définit de façon analogue une matrice *triangulaire inférieure.*

Des théorèmes 3 et 4 on déduit l'assertion suivante:

COROLLAIRE 1. — *Dans le cas où* $m=n$, *le système linéaire* (2) *est compatible et déterminé si, et seulement si, sa réduction à la forme quasi triangulaire donne un système* (5), *avec* $\bar{a}_{11}\bar{a}_{22} \ldots \bar{a}_{nn} \neq 0$. ▣

Il importe de noter que cette condition ne dépend pas des seconds membres du système. Aussi, pour $m = n$, le système (2) est-il compatible et déterminé si, et seulement si, cela est vrai pour le système homogène (2°) qui lui est associé. Or, un système homogène est

toujours compatible ; il a par exemple une solution nulle $x_1^\circ = 0, \ldots$
$\ldots, x_n^\circ = 0$.

La condition $\bar{a}_{11}\bar{a}_{22} \ldots \bar{a}_{nn} \neq 0$ signifie que le système homogène ne possède qu'une solution nulle. On peut donc donner au corollaire 1 un autre énoncé non lié à la forme quasi triangulaire :

COROLLAIRE 1'. — *Dans le cas où $m = n$, le système* (2) *est compatible et déterminé si, et seulement si, le système homogène* (2°) *qui lui est associé ne possède qu'une solution nulle.* ■

Il y a un cas qui mérite une attention spéciale, c'est celui où $n > m$.

COROLLAIRE 2. — *Lorsque $n > m$, le système compatible* (2) *est indéterminé. En particulier, le système homogène a toujours une solution non nulle quand $n > m$.*

En effet, dans tous les cas on a $r \leqslant m$, parce que le nombre d'équations dans le système (4) n'est pas plus grand que dans le système (2) (les équations, dont les premiers et seconds membres sont identiquement nuls, sont rejetées). L'inégalité $n > m$ entraîne donc $n > r$, ce qui signifie, en vertu du théorème 4, l'indétermination du système (2). Il reste à remarquer que l'indétermination d'un système homogène est équivalente à l'existence d'une solution non nulle pour ce système. ■

Une partie des résultats que nous avons obtenus est rassemblée dans le tableau ci-dessous.

	Type de système linéaire			
	général	homogène	$n > m$ non homogène	$n > m$ homogène
Nombre de solutions	0, 1, ∞	1, ∞	0, ∞	∞

5. **Quelques remarques et exemples.** — La méthode de résolution des systèmes d'équations linéaires que nous venons d'exposer, est connue sous le nom de *méthode de Gauss* ou *méthode d'éliminations successives des inconnues.* Cette méthode, bien commode pour de faibles valeurs de n, est aussi applicable à la résolution des problèmes sur calculateur électronique, bien qu'assez souvent et pour causes différentes, d'autres méthodes, par exemple les méthodes itératives, s'avèrent plus pratiques. Cela concerne surtout le cas où les coefficients sont donnés et les solutions sont cherchées avec un degré de précision imposé. Quant aux études théoriques, on y attache une importance primordiale à l'énoncé des conditions de compatibilité et de détermination d'un système linéaire, ainsi qu'à l'établisse-

ment des formules générales pour les résolutions en termes de coefficients et de termes constants sans réduire le système à la forme quasi triangulaire. Le corollaire 1′ répond, dans une certaine mesure, à l'une de ces exigences.

Exemple 1. — Revenons une fois de plus au problème de la plaquette chauffée du § 2. Comme nous l'avons vu au nº 1, la question qui nous intéresse, s'exprime par les propriétés d'un système linéaire parfaitement concret (appelons-le PC) à un assez grand nombre d'inconnues t_i. Suivant le critère énoncé dans le corollaire 1′, considérons un système linéaire homogène PCH associé à PC. En d'autres termes, la température de tous les points frontières de la plaquette est maintenant prise égale à zéro. Soit e le numéro d'un point intérieur ayant une valeur maximale de $|t_e|$. Alors, la condition

$$t_e = \frac{t_a + t_b + t_c + t_d}{4}$$

entraîne que $|t_e| = |t_a| = |t_b| = |t_c| = |t_d|$. En se déplaçant d'un pas de réseau dans n'importe lequel des quatre sens, on passera par des points ayant la même valeur $|t_i| = |t_e|$ tant qu'on n'atteindra un point frontière à température nulle. Cela signifie que $|t_e| = 0$ et donc $t_i = 0$ pour tous les i. Ainsi, le système PCH n'admet qu'une solution nulle, et PC est donc un système linéaire compatible et déterminé. Par là même, le problème de la plaquette chauffée, comme il a été initialement énoncé, est résolu.

Exemple 2. — Soit donné un système linéaire

$$\begin{aligned}
x_1 \quad \ldots\ldots\ldots\ldots &= 1,\\
x_2 \quad \ldots\ldots\ldots\ldots &= 1,\\
-x_1 - x_2 + x_3 \quad \ldots\ldots\ldots &= 0,\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots&\\
-x_{n-2} - x_{n-1} + x_n &= 0.
\end{aligned}$$

Il est évident que c'est un système compatible et déterminé, déjà réduit à la forme triangulaire. Seulement, pour sa résolution, il convient de se déplacer non pas de bas en haut, mais de haut en bas. En vertu de la définition même, sa solution est constituée par les n premiers *nombres de Fibonacci* $f_1, f_2, \ldots$ $\ldots, f_n$. Ces nombres se trouvent liés à un phénomène botanique appelé phyllotaxie (disposition des feuilles sur les plantes). Il serait bon cependant d'indiquer, pour $n = 1000$ et même pour n arbitraire, une expression générale (formule analytique) du n-ième nombre de Fibonacci. On pourrait objecter en disant qu'en suivant la définition, l'on aura assez de patience pour calculer même le nombre f_{1000}. Or, ce ne serait pas une résolution mathématique du problème. Aux chapitres 2 et 3 nous indiquerons deux expressions pour f_n, bien que ce problème concret puisse évidemment être résolu aussi par des méthodes plus directes.

Remarque. — Parfois, il s'avère plus commode de chercher les solutions d'un système linéaire sans le réduire à sa forme quasi triangulaire. Cela s'applique surtout au cas, où la matrice du système contient un grand nombre de zéros. Quelques exercices sont ici préférables à de longues explications.

§ 4. Déterminants d'ordre peu élevé

En exposant la méthode des éliminations successives des inconnues (méthode de Gauss) nous ne nous sommes pas trop préoccupés de coefficients des inconnues principales. Il nous importait seulement que ces coefficients soient différents de zéro. Nous allons maintenant appliquer cette méthode d'une façon plus correcte, tout au moins dans le cas des systèmes linéaires carrés d'équations à un nombre d'inconnues peu élevé. Ceci nous donnera matière à réflexion et une base de départ permettant de construire la théorie générale des déterminants que nous décrirons au chapitre 3.

Comme au § 3, considérons un système de deux équations à deux inconnues

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1, \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned} \tag{1}$$

et proposons-nous de chercher les formules générales pour les composantes x_1°, x_2° de sa solution. Appelons *déterminant* de la matrice

$$\left\| \begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{matrix} \right\|$$

l'expression $a_{11}a_{22} - a_{21}a_{12}$ et désignons-le par le symbole $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$.
Par là même, à la matrice carrée nous faisons correspondre un nombre

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}. \tag{2}$$

Cherchons à éliminer x_2 du système (1). A cet effet, multiplions la première équation par a_{22} et ajoutons-y la deuxième multipliée par $-a_{12}$. Nous obtenons

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} x_1 = b_1a_{22} - b_2a_{12}.$$

Le second membre de cette dernière égalité peut être considéré, lui aussi, comme un déterminant de la matrice

$\left\| \begin{matrix} b_1 & a_{12} \\ b_2 & a_{22} \end{matrix} \right\|$. Supposons que $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0$. Alors,

on a

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} \quad \text{et de façon analogue} \quad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}. \tag{3}$$

Connaissant les formules pour la résolution du système de deux équations linéaires à deux inconnues, nous pouvons résoudre certains

autres systèmes (résoudre des systèmes = trouver leurs solutions). Considérons, par exemple, un système de deux équations homogènes à trois inconnues:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= 0, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= 0. \end{aligned} \tag{4}$$

Nous nous intéressons à la solution non nulle de ce système, qui possède donc au moins l'un des $x_i \neq 0$. Soit , par exemple, $x_3 \neq 0$. En divisant les deux équations par $-x_3$ et en posant $y_1 = -x_1/x_3$, $y_2 = -x_2/x_3$, écrivons le système (4) sous la même forme

$$\begin{aligned} a_{11}y_1 + a_{12}y_2 &= a_{13}, \\ a_{21}y_1 + a_{22}y_2 &= a_{23}, \end{aligned}$$

que le système (1). Sous l'hypothèse que $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0$, les formules (3) donnent

$$y_1 = -\frac{x_1}{x_3} = \frac{\begin{vmatrix} a_{13} & a_{12} \\ a_{23} & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \quad y_2 = -\frac{x_2}{x_3} = \frac{\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}.$$

Ce n'est pas étonnant que nous avons déterminé les rapports des inconnues x_1, x_2, x_3 sans obtenir les inconnues elles-mêmes: du fait de l'homogénéité du système on déduit sans peine que si $(x_1^\circ, x_2^\circ, x_3^\circ)$ est solution et c un nombre arbitraire, alors $(cx_1^\circ, cx_2^\circ, cx_3^\circ)$ sera aussi solution du système. Par suite, nous pouvons poser

$$x_1 = -\begin{vmatrix} a_{13} & a_{12} \\ a_{23} & a_{22} \end{vmatrix}, \quad x_2 = -\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix}, \quad x_3 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \tag{5}$$

et dire que chaque solution est obtenue de la solution indiquée en multipliant tous les x_i par un certain nombre c. Pour donner à la réponse une forme plus symétrique, remarquons qu'on a toujours

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = -\begin{vmatrix} b & a \\ d & c \end{vmatrix},$$

ce qui résulte immédiatement de la formule (2). Aussi, les relations (5) peuvent-elles se mettre sous la forme

$$x_1 = \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}, \quad x_2 = -\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix}, \quad x_3 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}. \tag{6}$$

Ces formules sont obtenues à condition que $\begin{vmatrix} a_{11} & a_{21} \\ a_{21} & a_{22} \end{vmatrix} \neq 0$. On vérifie aisément que la proposition démontrée est vraie si l'un au moins des déterminants figurant dans les expressions (6) est différent de zéro. Dans le cas où tous les trois déterminants sont nuls,

les formules (6) donnent aussi, évidemment, une solution (à savoir, la solution nulle), mais on ne pourra plus affirmer que toutes les solutions du système sont obtenues sous forme d'un produit de la solution nulle par un nombre (considérer un système de deux équations identiques $x_1 + x_2 + x_3 = 0$).

Passons maintenant au cas d'un système de trois équations à trois inconnues

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3. \end{aligned}$$

Nous nous proposons d'éliminer x_2 et x_3 de ce système pour trouver la valeur de x_1. A cet effet, multiplions la première équation par c_1, la deuxième par c_2, la troisième par c_3 et additionnons les résultats. Choisissons les nombres c_1, c_2, c_3 de façon que dans l'équation obtenue après l'addition les termes en x_2 et x_3 deviennent égaux à 0. En annulant les coefficients correspondants, nous obtenons pour c_1, c_2, c_3 un système d'équations

$$\begin{aligned} a_{12}c_1 + a_{22}c_2 + a_{32}c_3 &= 0, \\ a_{13}c_1 + a_{23}c_2 + a_{33}c_3 &= 0, \end{aligned}$$

qui est du même type que (4). Par suite, on peut poser

$$c_1 = \begin{vmatrix} a_{22} & a_{32} \\ a_{23} & a_{33} \end{vmatrix}, \quad c_2 = -\begin{vmatrix} a_{12} & a_{32} \\ a_{13} & a_{33} \end{vmatrix}, \quad c_3 = \begin{vmatrix} a_{12} & a_{22} \\ a_{13} & a_{23} \end{vmatrix}.$$

Après avoir effectué des transformations évidentes, on obtient pour x_1 l'expression suivante:

$$\left(a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}\right)x_1 =$$
$$= b_1\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - b_2\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + b_3\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}. \quad (7)$$

Le coefficient de x_1 s'appelle déterminant de la matrice

$$\begin{Vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{Vmatrix} \text{ et se note } \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

Ainsi, pour déterminant du troisième ordre nous prenons l'expression

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} =$$
$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}, \quad (8)$$

définie à l'aide des déterminants du second ordre. On se rend compte sans peine que le second membre de l'égalité (7) s'obtient à partir du coefficient de x_1 en remplaçant a_{11} par b_1, a_{21} par b_2 et a_{31} par b_3. Il en découle que l'égalité (7) peut se mettre sous la forme

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} x_1 = \begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}.$$

Supposons que le coefficient de x_1 soit différent de zéro. Effectuant des calculs analogues pour x_2 et x_3, nous obtiendrons alors les formules:

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}, \quad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}, \quad x_3 = \frac{\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}. \qquad (9)$$

Il est évident que les mêmes raisonnements peuvent être appliqués à un système de quatre, cinq, etc. équations ayant le même nombre d'inconnues. A cet effet, nous devons d'abord obtenir des formules analogues à (6), qui donnent les solutions du système homogène de trois équations à quatre inconnues; puis, dans le système de quatre équations à quatre inconnues, éliminer x_2, x_3, x_4, en multipliant les équations par c_1, c_2, c_3, c_4 et en les additionnant. Nous trouverons les valeurs de c_i ($i = 1, 2, 3, 4$) à partir du système de trois équations homogènes.

Le coefficient de x_1, ainsi obtenu et construit à partir des déterminants du troisième ordre suivant le modèle (8), sera appelé déterminant du quatrième ordre. En développant les mêmes raisonnements pour x_2, x_3, x_4 nous obtiendrons pour x_i des formules analogues à (9). On pourra appliquer ce procédé un nombre indéfini de fois. La certitude que nous finirons par atteindre le but, nous est donnée par un principe général, largement utilisé en mathématiques, à savoir le principe d'induction mathématique ou principe de récurrence (voir plus loin § 7).

EXERCICES

1. La formule (8) devient plus facile à retenir si l'on utilise la règle des signes dont on affecte les produits qui entrent dans le développement d'un déterminant du troisième ordre (fig. 4). Enoncer une règle analogue pour un déterminant du quatrième ordre.

2. Montrer que les six termes figurant dans le développement d'un déterminant du troisième ordre ne peuvent pas être tous positifs.

3. Le carré de l'aire d'un parallélogramme construit sur les rayons vecteurs des points P, Q de coordonnées rectangulaires (α, β) et (γ, δ) (fig. 5) est donné

par la formule

$$\Delta^2=\begin{vmatrix}\alpha^2+\beta^2 & \alpha\gamma+\beta\delta\\ \alpha\gamma+\beta\delta & \gamma^2+\delta^2\end{vmatrix}.$$

(Il est bien facile de s'en assurer si l'on choisit un système de coordonnées dans lequel le point P se situe sur l'axe Ox.) Obtenir une formule analogue pour le

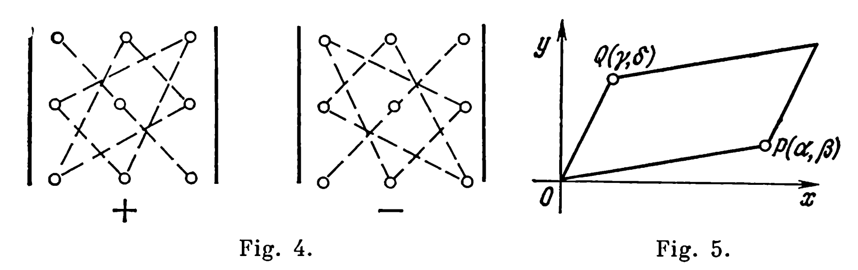


Fig. 4. Fig. 5.

carré du volume d'un parallélépipède dans l'espace à trois dimensions, en utilisant un déterminant du troisième ordre.

§ 5. Ensembles et applications

Dans les deux paragraphes qui précèdent, nous avons rencontré des ensembles d'éléments de nature différente, ainsi que des applications d'un ensemble dans l'autre. L'ensemble des solutions d'un système d'équations linéaires donné ou la relation qui à chaque matrice du deuxième ordre fait correspondre son déterminant, ce ne sont que des manifestations particulières des notions formelles dont la connaissance, ne serait-ce qu'au niveau intuitif, sera utile pour l'exposé ultérieur.

1. Ensembles.— On entend par *ensemble* toute collection ou tout assemblage d'objets appelés *éléments* d'ensemble. Un ensemble constitué d'un nombre fini d'éléments distincts peut être défini en extension : par une énumération explicite de tous ses éléments qu'on met généralement entre accolades. Ainsi, $\{1, 2, 4, 8\}$ désigne l'ensemble des puissances du nombre 2, comprises entre 1 et 10. En règle générale, on note l'ensemble par une lettre majuscule, et ses éléments, par des lettres minuscules du même ou d'un autre alphabet. Pour certains ensembles, particulièrement importants, on a adopté des notations canoniques qu'il convient de respecter. Ainsi, les lettres $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ désignent respectivement l'ensemble des entiers positifs (des entiers naturels), l'ensemble des entiers relatifs, l'ensemble des nombres rationnels et l'ensemble des nombres réels. Etant donné un ensemble S, la notation d'*appartenance* $a \in S$ indique que a appartient à S ou a est un élément de S ; dans le cas contraire on écrit $a \notin S$. On dit que S est une partie ou encore un *sous-ensemble* d'un ensemble T et on note $S \subset T$ (S est contenu ou inclus dans T) s'il

y a l'implication

$$x \in S, \ \forall x \Rightarrow x \in T.$$

(Pour les symboles, voir « Avis au lecteur », p. 14.) Deux ensembles coïncident (on dit encore : sont identiques ou égaux) si tout élément de chacun d'eux appartient à l'autre, ce qui se note

$$S = T \Leftrightarrow S \subset T \text{ et } T \subset S$$

($\Leftrightarrow$ se lit « si et seulement si » ou encore « équivalent à »). L'*ensemble vide* $\varnothing$, c'est-à-dire ne contenant aucun élément, est, par définition, une partie de tout ensemble. Si $S \subset T$ mais $S \neq \varnothing$ et $S \neq T$, on dit que S est une *partie propre* de T. Un sous-ensemble $S \subset T$ se définit souvent par une propriété quelconque que seuls les éléments de S possèdent. Ainsi, on définit par

$$\{ n \in \mathbb{Z} \mid n = 2m \text{ pour un } m \in \mathbb{Z}\}$$

l'ensemble de tous les entiers pairs, alors que

$$\mathbb{N} = \{n \in \mathbb{Z} \mid n > 0\}$$

désigne l'ensemble des entiers naturels.

On appelle *intersection* de deux ensembles S et T l'ensemble

$$S \cap T = \{x \mid x \in S \text{ et } x \in T\}.$$

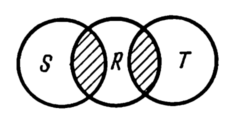
La *réunion* de deux ensembles S et T est par définition l'ensemble

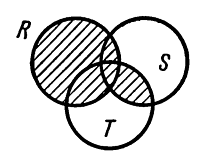
$$S \cup T = \{x \mid x \in S \text{ ou } x \in T\}.$$

L'intersection $S \cap T$ peut être un ensemble vide. Dans ce cas on dit que S et T sont des ensembles *disjoints*. Les opérations d'intersection et de réunion possèdent les propriétés suivantes

$$R \cap (S \cup T) = (R \cap S) \cup (R \cap T),$$
$$R \cup (S \cap T) = (R \cup S) \cap (R \cup T)$$

que le lecteur vérifiera à titre d'exercice. Les figures ci-dessous aideront à développer les raisonnements nécessaires, d'ailleurs assez simples :





On appelle *différence* $S \setminus T$ des ensembles S et T l'ensemble des éléments de S qui n'appartiennent pas à T. Ceci étant la condition $T \subset S$ n'est pas obligatoire en général. Au lieu de $S \setminus T$ on écrit aussi $S - T$.

Si T est une partie de S, la notation $S \setminus T$ désigne encore le *complémentaire* de T dans S. En posant $R = S \setminus T$, on aura: $R \cap T = \varnothing$, $R \cup T = S$. Le lecteur fera attention à une correspondance qui existe entre les notions d'intersection, de réunion et de complémentaire d'une part et les connecteurs logiques« ET »,« OU » et « NON » d'autre part.

Soient maintenant X et Y deux ensembles quelconques. On dit qu'un *couple* (x, y) d'éléments $x \in X$, $y \in Y$ pris dans cet ordre est un couple ordonné, et on considère que $(x_1, y_1) = (x_2, y_2)$ si, et seulement si, $x_1 = x_2$, $y_1 = y_2$. On appelle *produit cartésien* de deux ensembles X et Y l'ensemble de tous les couples (x, y):

$$X \times Y = \{(x, y) \mid x \in X, \ y \in Y\}.$$

Soit, par exemple, l'ensemble $\mathbb{R}$ de tous les nombres réels. Le *carré cartésien* $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ représente alors l'ensemble des toutes les coordonnées cartésiennes de tous les points du plan par rapport à un repère donné. De façon analogue, on introduit le produit cartésien $X_1 \times X_2 \times X_3$ de trois ensembles $(= (X_1 \times X_2) \times X_3 = X_1 \times \times (X_2 \times X_3))$, de quatre ensembles, etc. Pour $X_1 = X_2 = \ldots \ldots = X_k$ on écrit tout court $X^k = X \times X \times \ldots \times X$ et on dit que l'on a affaire à la *puissance cartésienne k-ième* de l'ensemble X. Les éléments de X^k sont des suites $(x_1, x_2, \ldots, x_k)$ de longueur k.

Pour mettre en évidence la différence qui existe entre les ensembles $X \times Y$ et $X \cup Y$, prenons pour X et Y des ensembles de puissance finie ou comme on le dit plus souvent de *cardinal* fini. Soit

$$|X| = \mathrm{Card}\, X = n, \qquad |Y| = \mathrm{Card}\, Y = m.$$

Alors

$$|X \times Y| = nm, \text{ et } |X \cup Y| = n + m - |X \cap Y|.$$

Si cela n'est pas clair, il convient de relire toutes les définitions.

2. Applications. — La notion d'*application* (de *fonction*) joue en mathématiques un rôle d'importance capitale. Etant donné les ensembles X et Y, l'application f dont le *domaine de définition* est X et le *domaine des valeurs* est Y fait correspondre à tout élément $x \in X$ un élément $f(x) \in Y$ que l'on désigne aussi par le symbole fx ou f_x. Dans le cas où $Y = X$, on dit que l'on a affaire à une application f de l'ensemble X dans lui-même. Symboliquement, l'application f s'écrit sous la forme $f\colon X \to Y$ ou $X \xrightarrow{f} Y$. On appelle *image* de l'application f l'ensemble de tous les éléments $f(x)$, tel que

$$\mathrm{Im}\, f = \{f(x) \mid x \in X\} = f(X) \subset Y.$$

L'ensemble

$$f^{-1}(y) = \{x \in X \mid f(x) = y\}$$

s'appelle *image réciproque* de l'élément $y \in Y$. Plus généralement, pour $Y_0 \subset Y$, posons

$$f^{-1}(Y_0) = \{x \in X \mid f(x) \in Y_0\} = \bigcup_{y \in Y_0} f^{-1}(y).$$

Si $y \in Y \setminus \operatorname{Im} f$, il est évident que $f^{-1}(y) = \varnothing$.

Une application $f\colon X \to Y$ s'appelle *application surjective* ou *surjection*, si $\operatorname{Im} f = Y$; elle est dite *injective* si $x \neq x'$ implique $f(x) \neq f(x')$. Enfin, une application $f\colon X \to Y$ est appelée *application bijective* ou *bijection* si elle est à la fois surjective et injective.

L'égalité de deux applications $f = g$ signifie par définition que leurs domaines correspondants coïncident : $X \xrightarrow{f} Y$, $X \xrightarrow{g} Y$ et $f(x) = g(x)$, $\forall x \in X$. Pour désigner une correspondance qui à un « argument » x, c'est-à-dire à un élément $x \in X$, associe un élément $f(x) \in Y$, on emploie la notation avec une flèche spéciale: $x \mapsto f(x)$ (se lit $f(x)$ est l'image de x par f).

Soit, par exemple, f_n un nombre de Fibonacci de numéro n (voir § 4). La correspondance $n \mapsto f_n$ définit une application $\mathbb{N} \to \mathbb{N}$ qui n'est ni surjective, ce qui est évident, ni injective parce que $f_1 = f_2 = 1$. Si $\mathbb{R}_+$ est l'ensemble de tous les nombres réels positifs, les applications $f\colon \mathbb{R} \to \mathbb{R}$, $g\colon \mathbb{R} \to \mathbb{R}_+$, $h\colon \mathbb{R}_+ \to \mathbb{R}_+$ définies par une seule et même loi $x \mapsto x^2$ sont toutes différentes: f n'est ni surjective, ni injective, g est surjective mais non injective, alors que l'application h est bijective. Ainsi, la donnée du domaine de définition et du domaine des valeurs est une partie essentielle de la définition d'une application (d'une fonction).

L'application $e_X\colon X \to X$ qui à tout élément $x \in X$ fait correspondre l'élément x lui-même s'appelle *application identique*. Si X est un sous-ensemble de Y: $X \subset Y$, il est parfois utile de considérer une application spéciale appelée *plongement* ou *immersion* $I\colon X \to Y$ qui à tout élément $x \in X$ associe le même élément, mais dans l'ensemble Y.

L'application $f\colon X \to Y$ s'appelle *restriction* (ou *contraction*) de l'application $g\colon X' \to Y'$ si $X \subset X'$, $Y \subset Y'$ et $f(x) = g(x)$, $\forall x \in X$. A son tour, g s'appelle *prolongement* de l'application f. Par exemple, le plongement $I\colon X \to Y$ est une restriction de l'application identique $e_Y\colon Y \to Y$.

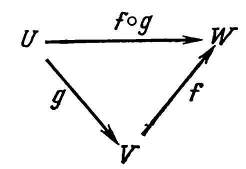
Nous aurons aussi à parler par la suite des fonctions de plusieurs variables. Il importe de saisir que la notion de puissance cartésienne X^n de l'ensemble X introduite plus haut permet de considérer la fonction $f(x_1, \ldots, x_n)$ de plusieurs variables $x_i \in X$, $i = 1, \ldots, n$, comme une application ordinaire $f\colon X^n \to Y$.

On appelle *composée* (ou *produit de composition*) de deux applications $g\colon U \to V$ et $f\colon V \to W$ une application notée $f \circ g\colon U \to W$

et définie par la condition

$$(f\circ g)(u) = f(g\,(u)),\ \forall u \in U.$$

Cette définition est illustrée par le *diagramme triangulaire*



On dit que ce diagramme « commute » (ou est *commutatif*), ce qui signifie que le résultat du passage de U à W est le même qu'on le fasse directement à l'aide de $f\circ g$ ou qu'on utilise une étape intermédiaire V. Remarquons que l'application composée n'est pas définie pour toutes les applications f et g. Il faut que dans leurs notations précédentes ces applications aient un ensemble commun V. Cependant la composée de deux applications d'un ensemble X dans lui-même a toujours un sens.

Par la suite, au lieu de $f\circ g$ nous écrirons tout simplement fg. Il est clair que

$$fe_X = f,\quad e_Yf = f$$

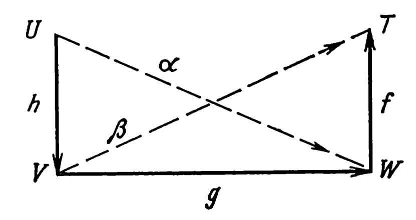
pour toute application $f\colon X \to Y$. La vérification de cette propriété est immédiate.

Une propriété importante de l'application composée est énoncée dans le théorème suivant.

Théorème 1. — *L'application composée obéit à la loi d'associativité. Ceci signifie que si* $h\colon U \to V$, $g\colon V \to W$, $f\colon W \to T$ *sont trois applications, on a*

$$f\,(gh) = (fg)\,h.$$

Démonstration. — Tous les raisonnements nécessaires sont mis en évidence par le diagramme



où $\alpha = gh$, $\beta = fg$. Conformément à la définition formelle de l'égalité des applications, il faut tout simplement comparer les valeurs des applications $f\,(gh)\colon U \to T$ et $(fg)\,h\colon U \to T$ en un « point » arbitraire $u \in U$. Or, de par la définition de l'application composée,

on a

$$(f\,(gh))\,u = f\,((gh)\,u\,) = f\,(g\,(hu)) = (fg)\,(hu) = ((fg)\,h)\,u. \blacksquare$$

La composition des applications $X \to X$ est en général non commutative, c'est-à-dire $fg \neq gf$. Il n'est pas difficile de s'en convaincre en considérant un exemple où $X = \{a, b\}$ est un ensemble à deux éléments, $f(a) = b$, $f(b) = a$, $g(a) = a$, $g(b) = a$. Un autre exemple : f et g sont deux applications *constantes* de X dans X, c'est-à-dire les valeurs $f(x)$ et $g(x)$ sont indépendantes de x. Alors on a $f \neq g \Rightarrow fg \neq gf$.

Certaines fonctions possèdent des *inverses*. Soient $f: X \to Y$ et $g: Y \to X$ deux applications quelconques dont les composées fg et gf sont définies. On dit que f est l'*inverse à gauche* de g et g l'*inverse à droite* de f si $fg = e_Y$. Dans le cas où les composées de f et g sont des applications identiques:

$$fg = e_Y, \quad gf = e_X, \tag{1}$$

on dit que g est *application inverse bilatère* (ou tout simplement *application inverse* ou *réciproque*) de f (et f est application réciproque de g) et on la note f^{-1}. Ainsi, $f(u) = v \Leftrightarrow f^{-1}(v) = u$.

En supposant qu'il existe encore une application $g': Y \to X$ pour laquelle

$$fg' = e_Y, \quad g'f = e_X, \tag{1'}$$

et en s'appuyant sur les égalités (1), (1') et le théorème 1, on obtient

$$g' = e_X g' = (gf)\,g' = g\,(fg') = ge_Y = g.$$

Ainsi, l'application réciproque de f, si elle existe, est unique, ce qui justifie l'emploi du symbole f^{-1}.

THÉORÈME 2. — *Une application $f: X \to Y$ admet une application réciproque si, et seulement si, elle est biunivoque (bijective).*

DÉMONSTRATION.— Pour démontrer le théorème on s'appuie sur une propriété qui, vu son importance, est présentée sous la forme du

LEMME. — *Si $f: X \to Y$, $g: Y \to X$ sont deux applications quelconques telles que $gf = e_X$, alors f est injective et g est surjective.*

En effet, soient $x, x' \in X$ et $f(x) = f(x')$. Alors, $x = e_X(x) = (gf)\,x = g\,(fx) = g\,(fx') = (gf)\,x' = e_X(x') = x'$. Par conséquent, f est injective. Si, ensuite, x est un élément quelconque de X, on a $x = e_X(x) = (gf)\,x = g\,(fx)$, ce qui prouve la surjectivité de g.

En revenant au théorème 2, supposons d'abord que f admette une application réciproque $g = f^{-1}$. Alors les égalités (1) et le lemme impliquent que f est à la fois surjective et injective, ce qui signifie qu'elle est une bijection. Réciproquement, en supposant f bijective,

nous trouverons pour tout $y \in Y$ un élément *unique* $x \in X$ pour lequel $f(x) = y$. En posant $g(y) = x$, nous définirons une application $g: Y \to X$ vérifiant les propriétés (1). Donc, $f^{-1} = g$. ■

COROLLAIRE.— *Si une application* $f: X \to Y$ *est bijective*, f^{-1} *l'est aussi et*

$$(f^{-1})^{-1} = f. \tag{2}$$

Soient de plus $f: X \to Y$, $h: Y \to Z$ *deux bijections. Alors leur composée* hf *est bijective et*

$$(hf)^{-1} = f^{-1}h^{-1}. \tag{3}$$

DÉMONSTRATION.— Si f est bijective, il existe, d'après le théorème 2, une application réciproque f^{-1}, ce qui signifie, en vertu du même théorème, que f^{-1} est bijective. La symétrie des conditions (1) écrites sous la forme $ff^{-1} = e_Y$, $f^{-1}f = e_X$ conduit à l'égalité (2). En outre, conformément à l'énoncé et au théorème 2, il existe des applications $f^{-1}: Y \to X$, $h^{-1}: Z \to Y$ et leur composée $f^{-1}h^{-1}: Z \to X$. Les égalités

$$(hf)(f^{-1}h^{-1}) = ((hf)f^{-1})h^{-1} = (h(ff^{-1}))h^{-1} = hh^{-1} = e_Z,$$

$$(f^{-1}h^{-1})(hf) = f^{-1}(h^{-1}(hf)) = f^{-1}((h^{-1}h)f) = f^{-1}f = e_X$$

entraînent que $f^{-1}h^{-1}$ est une application réciproque de hf. ■

Une application $\sigma: \mathbb{N} \to \mathbb{N}$, définie par $\sigma(n) = n + 1$, est une injection qui n'est pas surjective, vu que le premier élément (l'unité) n'appartient pas à $\operatorname{Im} \sigma$. Il est intéressant de noter que pour des ensembles finis une situation pareille ne se présente pas.

THÉORÈME 3. — *Si* X *est un ensemble fini et l'application* $f: X \to X$ *est injective, elle est aussi bijective.*

DÉMONSTRATION. — Il suffit de montrer que f est surjective, c'est-à-dire que pour tout élément $x \in X$ il existe x' avec $f(x') = x$. Posons

$$f^k(x) = f(f \ldots (fx) \ldots) = f(f^{k-1}x), \quad k = 0, 1, 2, \ldots$$

L'ensemble X étant fini, cette suite doit comporter des éléments répétés. Supposons que $f^m(x) = f^n(x)$, $m > n$. Si $n > 0$, alors l'égalité $f(f^{m-1}x) = f(f^{n-1}x)$ et l'injectivité de f entraînent $f^{m-1}(x) = f^{n-1}(x)$. En réitérant un nombre de fois suffisant la réduction de f, nous arriverons à un élément $x' = f^{m-n-1}(x)$ ayant la propriété exigée : $f(x') = x$. ■

Il est facile de comprendre qu'une *application surjective d'un ensemble fini dans lui-même est aussi une bijection.*

Quelques mots sur la puissance d'un ensemble. On dit que deux ensembles X et Y ont même puissance si, et seulement si, il existe une application bijective $f: X \to Y$. Les ensembles ayant même puissance que $\mathbb{N}$ (ou $\mathbb{Z}$) sont dits dénombrables.

EXERCICES

1. Soient $\Omega = \{+, -, ++, +-, -+, --, +++, \ldots\}$ l'ensemble de toutes les suites finies des signes $+$ et $-$ et $f: \Omega \to \Omega$ une application qui à un élément $\omega = \omega_1\omega_2 \ldots \omega_n \in \Omega$ fait correspondre $\omega' = \omega_1\bar{\omega}_1\omega_2\bar{\omega}_2 \ldots \ldots \omega_n\bar{\omega}_n$, où $\bar{\omega}_k = -$ si $\omega_k = +$ et $\bar{\omega}_k = +$ si $\omega_k = -$. Montrer que dans $f(f\omega)$ tout segment de longueur > 4 contient $++$ ou $--$.

2. L'application $f: \mathbb{N} \to \mathbb{N}$, définie par la loi $n \mapsto n^2$, admet-elle une application inverse à droite? Indiquer deux applications inverses à gauche de f. (Indication:

$$g_1(n) = [\sqrt{n}], \qquad \text{partie entière;}$$

$$g_2(n) = \begin{cases} \sqrt{n} & \text{si } n \text{ est un carré;} \\ 1 & \text{sinon.)} \end{cases}$$

3. Soit $f: X \to Y$ une application et soient S, T deux sous-ensembles de X. Montrer que

$$f(S \cup T) = f(S) \cup f(T), \quad f(S \cap T) \subset f(S) \cap f(T).$$

Donner un exemple montrant qu'en général la dernière inclusion ne peut pas être remplacée par une égalité.

4. La notation $\mathcal{P}(S) = \{T \mid T \subset S\}$ désigne l'ensemble de toutes les parties d'un ensemble S. Si, par exemple, $S = \{s_1, s_2, \ldots, s_n\}$ est un ensemble fini à n éléments, $\mathcal{P}(S)$ contient un ensemble vide $\varnothing$, n ensembles $\{s_1\}, \{s_2\}, \ldots \ldots, \{s_n\}$ à un élément, $n(n-1)/2$ ensembles $\{s_i, s_j \mid 1 \leqslant i < j \leqslant n\}$ et ainsi de suite jusqu'à $T = S$. Quelle est la puissance de l'ensemble $\mathcal{P}(S)$?

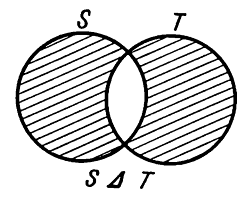
5. Soit $f: X \to Y$ une application et $b = f(a)$ pour un $a \in X$. L'image réciproque

$$f^{-1}(b) = f^{-1}(f(a)) = \{x \mid f(x) = f(a)\}$$

est encore appelée parfois *fibre* sur l'élément $b \in \operatorname{Im} f$. Montrer que l'ensemble X tout entier est réunion des fibres disjointes (partition de l'ensemble X). Avertissement: il convient de ne pas associer à la notation $f^{-1}(b)$ une application réciproque qui peut ne pas exister.

6. Montrer qu'une puissance cartésienne finie d'un ensemble dénombrable est un ensemble dénombrable.

7. Le symbole $S \triangle T$ désigne la *différence symétrique* de deux ensembles S et T:



$$S \triangle T = (S \setminus T) \cup (T \setminus S).$$

Montrer que

$$S \triangle T = (S \cup T) \setminus (S \cap T).$$

§ 6. Relations d'équivalence. Factorisation des applications

La notion d'équivalence des systèmes d'équations linéaires que nous avons introduite au § 3, incite à la considérer d'une façon générale, d'autant plus que nous utilisons les équivalences de divers types tant dans les raisonnements logiques que dans la vie courante.

1. Relations binaires.— Etant donné deux ensembles quelconques X et Y, une partie $O \subset X \times Y$ s'appelle *graphe de la relation binaire* de X vers Y (ou tout simplement relation binaire dans X si $Y = X$). Pour désigner qu'un couple (x, y) appartient à O, on utilise aussi la notation xOy et on dit que x est en relation O avec y. Une telle notation est bien commode, car elle permet d'éviter, par exemple, une expression encombrante

$$(x, y) \in O \ (O = <)$$

qui se remplace par une inégalité ordinaire $x < y$. Ici O est une relation d'ordre « $<$ » dans l'ensemble $\mathbb{R}$ des nombres réels, qui est

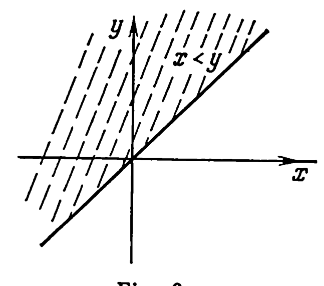


Fig. 6.

une relation binaire sur $\mathbb{R}$, formée de tous les points du plan $\mathbb{R}^2$ qui se situent au-dessus de la droite $x - y = 0$ (voir fig. 6).

A toute fonction $f: X \to Y$ on fait correspondre son graphe, c'est-à-dire le sous-ensemble

$$\Gamma(f) = \{(x, y) \mid x \in X, \ y = f(x)\} \subset X \times Y.$$

L'étude, dans $\mathbb{R}^2$, des graphes des fonctions $\mathbb{R} \to \mathbb{R}$ est du ressort du cours d'analyse mathématique. On comprend que tous les graphes O ne sont pas ceux des applications $X \to Y$. Une condition nécessaire et suffisante en est qu'à tout $x \in X$ corresponde exactement un seul élément y avec xOy. Au fait, la fonction f se trouve bien définie par X, Y et par son graphe $\Gamma(f)$.

2. Relation d'équivalence.— Une relation binaire $\sim$ dans X est appelée *relation d'équivalence* si tous les x, x', $x'' \in X$ vérifient

les conditions:

(i) $x \sim x$ (*réflexivité*);
(ii) $x \sim x' \Rightarrow x' \sim x$ (*symétrie*);
(iii) $x \sim x',\ x' \sim x'' \Rightarrow x \sim x''$ (*transitivité*).

La notation $a \nsim b$ signifie la négation de la relation d'équivalence entre les éléments $a,\ b \in X$.

Le sous-ensemble

$$\bar{x} = \{x' \in X \mid x' \sim x\} \subset X$$

de tous les éléments qui sont équivalents à un élément donné x s'appelle *classe d'équivalence* contenant x.

Puisque $x \sim x$ (voir (i)), on a en effet $x \in \bar{x}$. Tout élément $x' \in \bar{x}$ s'appelle *représentant* de la classe $\bar{x}$. On a l'assertion suivante:

L'ensemble des classes d'équivalence par la relation $\sim$ est une partition de l'ensemble X (notée $\pi_\sim(X)$) *en ce sens que X est réunion des sous-ensembles disjoints.*

En effet, puisque $x \in \bar{x}$, on a $X = \bigcup\limits_{x \in X} \bar{x}$. La classe $\bar{x}$ est définie univoquement par l'un quelconque de ses éléments, c'est-à-dire $\bar{x} = \bar{x}' \Leftrightarrow x \sim x'$. Démontrons l'implication $x \sim x' \Rightarrow \bar{x} = \bar{x}'$. On a: $x \sim x'$ et $x'' \in \bar{x} \Rightarrow x'' \sim x \Rightarrow x'' \sim x' \Rightarrow x'' \in \bar{x}' \Rightarrow \bar{x} \subset \bar{x}'$. Mais $x \sim x' \Rightarrow x' \sim x$ (voir (ii)), ce qui veut dire que l'inclusion réciproque $\bar{x}' \subset \bar{x}$ est aussi vraie. Donc, $\bar{x}' = \bar{x}$. Réciproquement: puisque $x \in \bar{x}$, on a $\bar{x}' = \bar{x} \Rightarrow x \in \bar{x}' \Rightarrow x \sim x'$.

Si, maintenant, $\bar{x}' \cap \bar{x}'' \neq \varnothing$ et $x \in \bar{x}' \cap \bar{x}''$, on a $x \sim x'$ et $x \sim x''$, d'où, en vertu de la transitivité (iii), on a $x' \sim x''$ et $\bar{x}' = \bar{x}''$. Cela signifie que les différentes classes sont disjointes. ∎

Soit $\Pi = \mathbb{R}^2$ un plan réel rapporté à un système de coordonnées rectangulaires. La relation $\sim$ dans Π, définie par l'appartenance des points $P,\ P' \in \Pi$

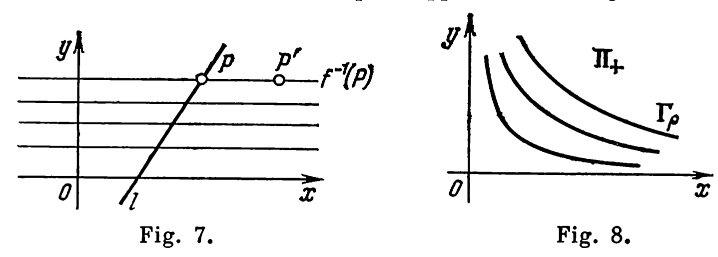


Fig. 7. Fig. 8.

à une même droite horizontale, est évidemment une relation d'équivalence dont les classes sont des droites horizontales (fig. 7). Les hyperboles Γ_ρ (fig. 8) de la forme $xy = \rho > 0$ définissent la relation d'équivalence dans le sous-ensemble $\Pi_+ \subset \Pi$ des points $P(x,\ y)$ de coordonnées $x > 0,\ y > 0$. Ces exemples géométriques permettent d'énoncer l'assertion réciproque suivante.

Etant donné une partition quelconque $\pi(X)$ *de l'ensemble* X *en sous-ensembles disjoints* C_x, *les parties* C_x *sont des classes d'équivalence par une certaine relation d'équivalence* $\sim$.

En effet, par hypothèse, tout élément $x \in X$ appartient exactement à un seul sous-ensemble C_a. Posons $x \sim x'$ si, et seulement si, x et x' appartiennent à un seul et même sous-ensemble C_a. Il est évident que la relation $\sim$ est réflexive, symétrique et transitive, donc $\sim$ est une relation d'équivalence. Il vient par définition de $\sim$ que $x \in C_a \Rightarrow \bar{x} = C_a$. Par suite, $\pi(X) = \pi_{\sim}(X)$. ■

3. Factorisation des applications.— Etant donné la correspondance biunivoque établie ci-dessus entre les relations d'équivalence et les partitions de l'ensemble X, on convient d'appeler *ensemble quotient* S par la relation d'équivalence $\sim$ et de le noter $X/\sim$ une partition correspondant à cette relation d'équivalence. L'application surjective

$$p : x \mapsto p(x) = \bar{x} \tag{1}$$

s'appelle *application canonique* de X sur l'ensemble quotient $X/\sim$.

Soient X, Y deux ensembles et $f: X \to Y$ une application. La relation binaire O_f:

$$xO_fx' \Leftrightarrow f(x) = f(x'), \quad \forall x, x' \in X,$$

est manifestement réflexive $(f(x) = f(x))$, symétrique $(f(x') = f(x) \Rightarrow f(x) = f(x'))$ et transitive $(f(x) = f(x')$ et $f(x') = f(x'') \Rightarrow f(x) = f(x''))$. Ainsi, O_f est une relation d'équivalence dans X. Les classes d'équivalence correspondantes $\bar{x}$ sont des fibres (images réciproques) au sens de l'exercice 5 du § 5. En d'autres termes

$$\bar{x} = \{x' \mid f(x') = f(x)\}.$$

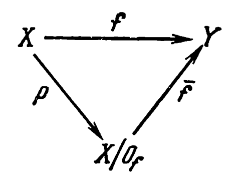
L'application $f: X \to Y$ *induit* une application $\bar{f}: X/O_f \to Y$ définie par

$$\bar{f}(\bar{x}) = f(x), \tag{2}$$

ou, ce qui revient au même,

$$\bar{f}p(x) = f(x), \tag{2'}$$

où p est une application canonique (1). Puisque $\bar{x} = \bar{x}' \Leftrightarrow f(x) = f(x')$, la relation (2) qui définit $\bar{f}$, ne dépend pas du représentant x choisi dans la classe $\bar{x}$. Dans de tels cas on dit que la définition de $\bar{f}$ est *correcte*. Le diagramme commutatif



illustre la *factorisation* (*décomposition*)

$$f = \bar{f} \cdot p \tag{3}$$

de l'application f en un produit de la surjection p et de l'injection $\bar{f}$. L'injectivité de $\bar{f}$ découle de ce que

$$\bar{f}(\bar{x}_1) = \bar{f}(\bar{x}_2) \Leftrightarrow f(x_1) = f(x_2) \Leftrightarrow \bar{x}_1 = \bar{x}_2.$$

La bijectivité de $\bar{f}$ est équivalente à la surjectivité de f. Remarquons que si $f' : X/O_f \to Y$ est une autre application pour laquelle est vérifiée la relation (3): $f'p = f$, alors $f'(\bar{x}) = f'(px) = (f'p)x = f(x) = \bar{f}(\bar{x})$ (voir (2)) implique l'égalité $f' = \bar{f}$. Il en résulte l'unicité de l'application $\bar{f}$ qui rend commutatif le diagramme triangulaire indiqué ci-dessus.

4. Ensembles ordonnés.— Une relation binaire $\leqslant$ dans un ensemble X est appelée *relation d'ordre sur* X si elle est réflexive ($x \leqslant x$), antisymétrique ($x \leqslant y$ et $y \leqslant x$ entraînent $x = y$) et transitive ($x \leqslant y$ et $y \leqslant z$ entraînent $x \leqslant z$). Pour $x \leqslant y$ et $x \neq y$ on écrit $x < y$. Au lieu de $x \leqslant y$ on utilise aussi la notation $y \geqslant x$. Deux éléments $x, x' \in X$ peuvent ne pas être en relation $\leqslant$. Si, pourtant, $x \leqslant x'$ ou $x' \leqslant x$ pour tout couple d'éléments de X, on dit que X est un ensemble *totalement ordonné* (ou une *chaîne*). Dans le cas contraire, l'*ordre* sur X est qualifié de *partiel*.

L'ensemble $X = \mathscr{P}(S)$ des parties d'un ensemble S (voir exercice 4 du § 5) muni de la relation d'inclusion ordinaire $R \subset T$ entre les sous-ensembles, et l'ensemble $\mathbb{N}$ des entiers naturels muni de la relation $d \mid n$ (n est divisible par d) sont des ensembles partiellement ordonnés.

Soit X un ensemble partiellement ordonné quelconque et soient x et y ses éléments. On dit que y *revêt* x si $x < y$ et il n'existe pas de z tel que $x < z < y$. Dans le cas où Card $X < \infty$, $x < y$ (c'est-à-dire x et y sont comparables) si, et seulement si, il existe une suite d'éléments $x = x_1, x_2, \ldots, x_{n-1}, x_n = y$ dans laquelle x_{i+1} revêt x_i. La notion de revêtement est commode pour représenter un ensemble fini partiellement ordonné X par un diagramme plan. Les éléments de l'ensemble X sont représentés par des points. Si y revêt x, on le situe plus haut que x et on joint x à y par un segment de droite. La comparabilité de y et de x se représente par une ligne brisée « descendante » qui joint x et y, le nombre de telles lignes étant parfois supérieur à un. On ne trace aucune ligne si x et y ne sont pas comparables. Deux des trois diagrammes de la fig. 9 représentent un « segment » de la suite naturelle $\mathbb{N}$ totalement ordonnée et un ensemble $\mathscr{P}(\{a, b, c\})$ avec la relation d'ordre sur $\mathscr{P}(S)$ introduite plus haut.

On appelle *le plus grand élément* d'un ensemble X partiellement ordonné un élément $n \in X$, tel que $x \leqslant n$ pour tous les $x \in X$. Un élément $m \in X$ s'appelle *élément maximal* si $m \leqslant x \in X$ implique $x = m$. Le plus grand élément est toujours maximal. La réciproque est fausse. Les éléments maximaux peuvent être nombreux, alors que le plus grand élément, s'il existe, est unique. Les mêmes remarques

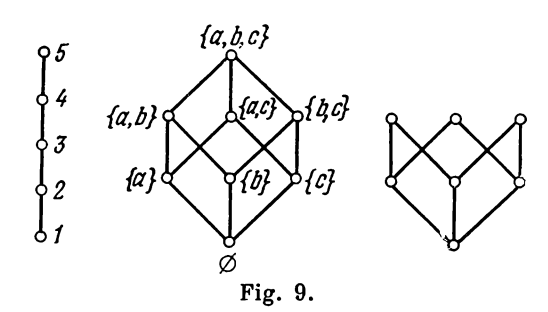


Fig. 9.

s'appliquent au *plus petit élément* et à l'*élément minimal*. Deux diagrammes de la fig. 9, à gauche, comportent les plus grands et les plus petits éléments. Sur le diagramme à droite on voit trois éléments maximaux, un élément qui est le plus petit et aucun qui soit le plus grand.

Riche de nombreux résultats bien importants, la théorie des systèmes algébriques partiellement ordonnés (algèbres de Boole, treillis) occupe une place considérable en algèbre. Or, l'étude de ces systèmes sortirait nettement du cadre du présent ouvrage. Ce paragraphe avait un objectif plus modeste de familiariser le lecteur avec une autre relation binaire et de lui donner une idée des diagrammes qui permettront par la suite de mieux comprendre la disposition relative des sous-groupes dans les groupes ou, par exemple, la disposition des sous-corps dans les corps.

EXERCICES

1. Montrer qu'il existe une bijection entre l'ensemble quotient $\mathbb{R}^2/\sim$ obtenu à partir de la représentation géométrique de la fig. 7 et toute droite l qui coupe l'axe Ox.

2. Poser $P(x, y) \sim P(x', y')$ pour les points du plan $\mathbb{R}^2$ si, et seulement si, $x' - x \in \mathbb{Z}$ et $y' - y \in \mathbb{Z}$. Démontrer que $\sim$ est une relation d'équivalence et que l'ensemble quotient $\mathbb{R}^2/\sim$ est représenté géométriquement par les points sur un tore (voir fig. 10).

3. Montrer que les ensembles à deux, trois et quatre éléments possèdent respectivement 2, 5 et 15 ensembles quotients différents.

4. Soit $\sim$ une relation d'équivalence dans un ensemble X et soit $f: X \to Y$ une application pour laquelle $x \sim x' \Rightarrow f(x) = f(x')$. Montrer que cette condition de *compatibilité* de f avec $\sim$ (plus faible que celle examinée au n° 2) permet de définir correctement l'application induite $\bar{f}: \bar{x} \mapsto f(x)$ de $X/\sim$ dans Y, qui

conduit à la factorisation $f = \bar{f} \cdot p$, $\bar{f}$ n'étant plus nécessairement injective. Quelle est la condition d'injectivité de $\bar{f}$?

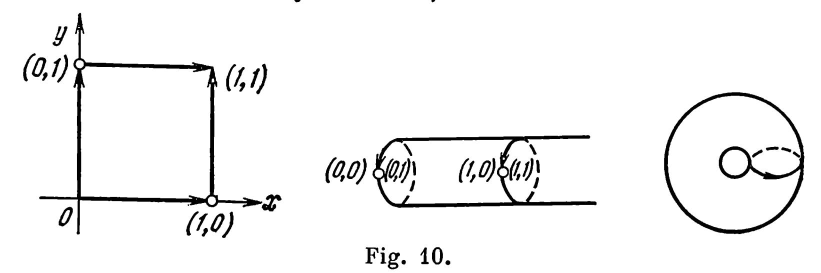


Fig. 10.

5. Représenter par des diagrammes les ensembles partiellement ordonnés: 1) $\mathcal{P}(\{a, b, c, d\})$; 2) l'ensemble de tous les diviseurs de 24 (les relations d'ordre sont définies dans le texte).

§ 7. Principe d'induction mathématique (de récurrence)

On suppose connu l'ensemble $\mathbb{N} = \{1, 2, 3, \ldots\}$ de tous les *nombres naturels* (c'est-à-dire des *entiers positifs*). Or, en réalité ce sont les axiomes de Peano (1858-1932) qui ont été placés au point de départ de l'étude de $\mathbb{N}$. De ses axiomes (que nous ne reproduisons pas) on déduit les propriétés de l'addition, de la multiplication et de l'ordre total (voir § 6, n° 4) des entiers naturels ou plus exactement du système $\mathbb{N} \cup \{0\}$. En particulier, on démontre une proposition intuitivement claire: *tout sous-ensemble non vide $S \subset \mathbb{N}$ possède un plus petit élément*, c'est-à-dire un entier naturel $s \in S$ qui est inférieur à tous les autres nombres de S. Compte tenu de cette assertion, on déduit des axiomes de Peano le principe suivant.

Principe de récurrence. — *Supposons que nous ayons une proposition $M(n)$ relative à tout $n \in \mathbb{N}$. Supposons aussi que nous disposions d'une règle permettant d'établir la vérité de $M(l)$ pour un l donné, à condition que $M(k)$ soit vraie pour tout $k < l$ (on suppose en particulier que nous pouvons vérifier la vérité de M (1)). Alors, $M(n)$ est vraie pour tout $n \in \mathbb{N}$.*

Raisonnons par l'absurde: supposons que le sous-ensemble

$$S = \{s \mid s \in \mathbb{N},\ M(s) \text{ est fausse}\} \subset \mathbb{N}$$

ne soit pas vide. D'après ce qui précède, S aurait un plus petit élément s_0. Alors, la proposition $M(s_0)$ serait fausse et $M(s)$ vraie pour tout $s < s_0$. Pourtant, cela contredit l'hypothèse de pouvoir démontrer la vérité de $M(s_0)$. ∎

Il ne s'agit pas ici d'étudier d'une manière détaillée le principe d'induction mathématique. Nous nous contenterons d'indiquer que ce principe reflète, si l'on peut dire ainsi, l'essentiel de la suite natu-

relle, et que la connaissance de celle-ci ne se réduit pas à quelque chose notablement plus simple.

Il importe encore de signaler un point obligatoire que comporte la « démonstration par induction complète » ou, comme on le dit le plus souvent, la démonstration par récurrence. Il consiste à établir la *base de récurrence*, c'est-à-dire à vérifier que la propriété ou la proposition donnée sont vraies pour de petites valeurs de n. Sans une telle vérification on pourrait arriver à des conclusions spéculatives arbitraires du type « tous les étudiants sont de même taille ». Le raisonnement qui peut conduire à une telle conclusion est le suivant : l'ensemble vide des étudiants et l'ensemble à un seul étudiant vérifient cette propriété. Avançons l'hypothèse de récurrence que tout ensemble de $\leqslant n$ étudiants possède cette propriété. Dans l'ensemble des $n+1$ étudiants, les n premiers et les n derniers étudiants sont de même taille par l'hypothèse de récurrence. L'intersection de ces ensembles est un sous-ensemble des $n-1$ étudiants toujours de même taille. Par conséquent, tous les $n+1$ étudiants sont de même taille. En réalité, la première proposition à étudier devrait se rapporter à un ensemble de deux étudiants quelconques. Or, c'est elle qui est ici fausse. Quelle doit donc être la longueur de la base de récurrence? Généralement, on la détermine selon la démonstration. Dans notre cas élémentaire, une condition importante est que l'intersection de deux ensemble ne soit pas vide, c'est-à-dire que soit vérifiée l'inégalité $n-1 \geqslant 1$, d'où $n \geqslant 2$.

Dans des situations plus complexes et en particulier dans les cas où il s'agit de définir ou de construire un être mathématique par induction, à l'aide de relations récurrentes (comme nous le ferons au chapitre 3 pour les déterminants des matrices), il doit être apporté une attention spéciale au choix de la base de récurrence. D'autre part, il ne faut pas sacrifier à un autre excès : après s'être assuré de la vérité de $M(k)$ pour tous les k appartenant à un segment suffisamment grand $1 \leqslant k \leqslant l$ de la suite naturelle, faire une conclusion non fondée sur la vérité de $M(n)$ pour tous les $n \in \mathbb{N}$ (ce sera une induction incomplète).

En voici deux exemples.

1. P. Fermat supposait que tous les nombres de la forme $F_n = 2^{2^n}+1$, $n = 0, 1, \ldots$ (*nombres de Fermat*) étaient premiers. En effet, les cinq premiers nombres de Fermat sont premiers, mais pour F_5 Euler trouve la décomposition $F_5 = 4\,294\,967\,297 = 641 \cdot 6\,700\,417$. Les tentatives persévérantes faites en vue d'obtenir, à l'aide de calculateurs électroniques les plus modernes, au moins un nouveau nombre de Fermat qui soit premier, ne sont pas encore couronnées de succès. L'une des dernières « acquisitions » réalisées dans ce domaine est la vérification que F_{1945} est divisible par $5 \cdot 2^{1947}+1$.

2. L'étude, pour $n = 1, 2, \ldots, 40$, des nombres de la forme $n^2 - n + 41$ (ce polynôme a été proposé par Euler) peut inciter à penser que ces nombres sont premiers pour tout n (pour les nombres premiers voir § 8). Pourtant, $41^2 - 41 + 41 = 41^2$.

On peut fournir de tels exemples aussi nombreux que l'on voudrait.

Parfois, dans les raisonnements par récurrence, l'essentiel est de donner une forme adéquate à la proposition à démontrer. Soit à calculer la somme

$$p_k(n) = 1^k + 2^k + 3^k + \ldots + (n-1)^k + n^k, \qquad k = 1, 2, 3.$$

Le problème devient sensiblement plus facile à résoudre s'il est connu que la réponse supposée est contenue dans les expressions:

$$p_1(n) = \frac{n(n+1)}{2}, \quad p_2(n) = \frac{n(n+1)(2n+1)}{6},$$
$$p_3(n) = \left[\frac{n(n+1)}{2}\right]^2.$$

Si $p_1(n)$ n'est pas difficile à établir (cela a été fait par Gauss encore en jeune âge), la forme de $p_2(n)$ et $p_3(n)$ n'est pas si triviale, alors que la relation

$$p_5(n) + p_7(n) = 2\left[\frac{n(n+1)}{2}\right]^4$$

devrait être cherchée suivant un plan bien déterminé. Dans le cas considéré, un tel plan peut être indiqué, mais ici il ne s'agit pas de cela. Pour justifier toutes les relations indiquées ci-dessus, il faut effectuer, par des calculs directs, la récurrence de n à $n+1$. Nous conseillons au lecteur de le faire à titre d'exercice bien utile.

A propos, pour cet exercice, on aura besoin d'une formule dite *formule du binôme*:

$$(a+b)^n = a^n + \binom{n}{1}a^{n-1}b + \ldots + \binom{n}{k}a^{n-k}b^k + \ldots + b^n. \quad (1)$$

Ici a et b sont des nombres arbitraires, et le *coefficient binomial* $\binom{n}{k}$ du monôme $a^{n-k}b^k$ est de la forme

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!} = \frac{n(n-1)\ldots(n-k+1)}{k(k-1)\ldots 2\cdot 1}. \quad (2)$$

Il est utile de compléter (2) par la convention que $0! = 1$ et $\binom{n}{k} = 0$ pour $k < 0$. Remarquons encore que

$$\binom{n}{n-k} = \binom{n}{k}$$

(la propriété de symétrie des coefficients binomiaux).

Nous allons démontrer la formule (1), qui est manifestement vraie pour $n = 1, 2$, par récurrence sur n. En la supposant vraie pour tous les exposants $\leqslant n$, multiplions les deux membres de la relation (1)

par $a + b$. Il vient

$$(a+b)^{n+1} = (a+b)^n (a+b) =$$
$$= a^n (a+b) + \ldots + \binom{n}{k} a^{n-k} b^k (a+b) + \ldots + b^n (a+b) =$$
$$= a^{n+1} + a^n b + \ldots + \binom{n}{k-1} a^{n+2-k} b^{k-1} + \binom{n}{k-1} a^{n+1-k} b^k +$$
$$+ \binom{n}{k} a^{n+1-k} b^k + \binom{n}{k} a^{n-k} b^{k+1} + \ldots + ab^n + b^{n+1}.$$

La réduction des termes semblables montre que le coefficient du binôme $a^{n+1-k} b^k$ sera

$$\binom{n}{k-1} + \binom{n}{k} =$$
$$= \frac{n!}{(k-1)!\,(n-k+1)!} + \frac{n!}{k!\,(n-k)!} = \frac{n!}{(k-1)!\,(n-k)!} \left[\frac{1}{n-k+1} + \frac{1}{k} \right] =$$
$$= \frac{n!}{(k-1)!\,(n-k)!} \cdot \frac{n+1}{k\,(n-k+1)} = \frac{(n+1)!}{k!\,(n+1-k)!} = \binom{n+1}{k}.$$

C'est justement le coefficient binomial de la forme (2) à indice supérieur augmenté d'une unité. Par là même, la validité de la formule (1) est démontrée pour tout $n \in \mathbb{N}$. Soit

$$(a + b)^n = (a + b)\,(a + b)\,\ldots\,(a + b);$$

si l'on affecte chaque facteur du second membre d'un numéro de 1 à n, et qu'on considère ceux des sous-ensembles de numéros $1 \leqslant i_1 < i_2 < \ldots < i_k \leqslant n$, qui correspondent après la multiplication au monôme $a^{n-k} b^k$, il vient que $\binom{n}{k}$ *n'est rien d'autre que le nombre de toutes les parties de puissance* k *de l'ensemble à* n *éléments.* Le terme un peu démodé — le nombre $C_n^k = \binom{n}{k}$ de combinaisons de n éléments k à k — exprime au fond la même chose.

En particulier, la puissance de l'ensemble $\mathcal{P}(\{s_1, \ldots, s_n\})$ (voir exercice 4 du § 5) est égale à $\binom{n}{0} + \binom{n}{1} + \ldots + \binom{n}{n-1} + \binom{n}{n}$. Posant dans la formule (1) $a = b = 1$, on obtient

$$2^n = \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \ldots + \binom{n}{n-1} + \binom{n}{n}.$$

Ainsi, $\mathrm{Card}\, \mathcal{P}(\{s_1, s_2, \ldots, s_n\}) = 2^n$.

Parfois, il est très commode de démontrer un théorème ou de construire un être mathématique en se basant sur des formes plus complexes de récurrence. Considérons à titre d'exemple le principe d'« induction double ». On associe à deux entiers naturels quelconques m et n une assertion $A(m, n)$ telle que: (i) $A(m, 1)$ et $A(1, n)$ sont vraies pour tous les m et n; (ii) si $A(k - 1, l)$ et $A(k, l - 1)$ sont vraies, $A(k, l)$ est aussi vraie (équivalent: (ii)' si $A(k', l')$

est vraie pour tous les $k' \leqslant k$, $l' \leqslant l$, $k' + l' < k + l$, alors $A(k, l)$ est aussi vraie). Alors, l'assertion $A(m, n)$ est vraie quels que soient les entiers naturels m et n.

§ 8. Arithmétique des nombres entiers

Ce paragraphe a pour but de décrire succinctement les propriétés les plus simples de la divisibilité des nombres entiers auxquelles il sera commode par la suite de faire référence à différentes occasions. Des résultats supplémentaires seront indiqués au chapitre 5, où la théorie de la divisibilité sera étendue à des systèmes algébriques plus généraux.

1. Théorème fondamental de l'arithmétique.— Un entier s s'appelle *diviseur* (ou *facteur*) d'un nombre entier n si $n = st$ pour un certain $t \in \mathbb{Z}$. A son tour, n s'appelle *multiple* de s. La divisibilité de n par s se note $s \mid n$, et la négation de cette divisibilité, $s \nmid n$. La divisibilité est une relation transitive dans $\mathbb{Z}$. Si $m \mid n$ et $n \mid m$, alors $n = \pm m$, et les entiers n, m sont dits *associés*. Un nombre entier p qui n'admet pour diviseur que les nombres $\pm p$, ± 1 (*diviseurs non propres*) s'appelle *nombre premier*. Comme nombres premiers on prend généralement les nombres premiers positifs > 1.

Le rôle fondamental des nombres premiers est mis en évidence par le théorème suivant :

THÉORÈME FONDAMENTAL DE L'ARITHMÉTIQUE. — *Tout nombre entier positif $n \neq 1$ peut s'écrire sous la forme d'un produit de nombres premiers : $n = p_1 p_2 \ldots p_s$. Cette écriture est unique à un ordre des facteurs près.*

En groupant ensemble les facteurs premiers identiques et en modifiant les désignations, on obtient l'expression de n sous la forme $n = p_1^{\varepsilon_1} p_2^{\varepsilon_2} \ldots p_k^{\varepsilon_k}$, $\varepsilon_i > 0$, $1 \leqslant i \leqslant k$. Pour tout nombre rationnel $a = n/m \in \mathbb{Q}$, on a une décomposition analogue, mais avec des exposants ε_i tant positifs que négatifs.

Remarquons que l'*ensemble*

$$P = \{2, 3, 5, 7, 11, 13, \ldots\}$$

de tous les nombres premiers est infini (théorème d'Euclide). En effet, s'il n'existait qu'un ensemble fini de nombres premiers, disons $p_1, p_2, \ldots, p_t$, alors, en vertu du théorème fondamental, le nombre $c = p_1 p_2 \ldots p_t + 1$ serait divisible par au moins un nombre pris parmi les p_i. Posons, sans restreindre la généralité, $c = p_1 c'$. Alors, $p_1 (c' - p_2 \ldots p_t) = 1$, ce qui est impossible parce que, dans $\mathbb{Z}$, les seuls diviseurs de l'unité sont ± 1. ∎

La démonstration du théorème fondamental de l'arithmétique sera remise au chapitre 5. Il semble à première vue que ce théorème est tellement évident qu'il n'exige en général aucune démonstration. Pourtant, bien qu'il s'agisse

des propriétés multiplicatives (des propriétés de la divisibilité) des nombres entiers, il est impossible de démontrer le théorème fondamental sans utiliser à la fois les opérations de multiplication et d'addition dans $\mathbb{Z}$. Pour illustrer cette assertion, considérons dans $\mathbb{N}$ le sous-ensemble $S = \{4k + 1 \mid k = 0, 1, 2, \ldots\}$. Il est stable pour la multiplication : $(4k_1 + 1)(4k_2 + 1) = 4k_3 + 1$. En raisonnant par récurrence sur $n \in S$ on établit sans peine l'existence de la décomposition (première partie du théorème fondamental) $n = q_1 \ldots q_l$, q_i étant des éléments de S qui ne sont plus factorisables. Nous les appellerons nombres quasi premiers. Ecrivons quelques-uns de ces nombres: 5, 9, 13, 17, 21, 49. Quant à la deuxième partie du théorème fondamental, elle n'est pas vraie pour S, car le nombre $441 \in S$, par exemple, admet deux décompositions tout à fait différentes en un produit de nombres quasi premiers:

$$441 = 9 \cdot 49 = 21^2.$$

2. P.G.C.D. et P.P.C.M. dans $\mathbb{Z}$.— Deux entiers, quels qu'ils soient, n et m peuvent s'écrire sous la forme d'un produit de mêmes nombres premiers

$$n = \pm\, p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_k^{\alpha_k}, \quad m = \pm\, p_1^{\beta_1} p_2^{\beta_2} \ldots p_k^{\beta_k},$$

si l'on convient d'admettre des exposants nuls (en considérant comme toujours que $p_i^0 = 1$). Introduisons deux entiers en les définissant par les relations

$$\text{P.G.C.D.}\,(n, m) = p_1^{\gamma_1} p_2^{\gamma_2} \ldots p_k^{\gamma_k}, \quad \text{P.P.C.M.}\,(n, m) = p_1^{\delta_1} p_2^{\delta_2} \ldots p_k^{\delta_k}, \tag{1}$$

où $\gamma_i = \min(\alpha_i, \beta_i)$, $\delta_i = \max(\alpha_i, \beta_i)$, $i = 1, 2, \ldots, k$. Puisque $d \mid n \Rightarrow d = \pm p_1^{\alpha'_1} \ldots p_k^{\alpha'_k}$, $0 \leqslant \alpha'_i \leqslant \alpha_i$, les relations (1) entraînent les propositions suivantes:

(i) P.G.C.D. $(n, m) \mid n$, P.G.C.D.$(n, m) \mid m$ et si $d \mid n$, $d \mid m$, alors $d \mid$ P.G.C.D.(n, m);

(ii) $n \mid$ P.P.C.M.(n, m), $m \mid$ P.P.C.M.(n, m) et si $n \mid u$, $m \mid u$, alors P.P.C.M.$(n, m) \mid u$.

Les propriétés (i) et (ii) justifient les dénominations de plus grand commun diviseur (P.G.C.D) et de plus petit commun multiple (P.P.C.M.) des entiers n et m. Pour $n > 0$, $m > 0$ on a la relation

$$\text{P.G.C.D.}(n, m) \cdot \text{P.P.C.M.}(n, m) = nm. \tag{2}$$

On dit que n et m sont des *nombres premiers entre eux* si P.G.C.D. $(n, m) = 1$. Dans ce cas, la relation (2) prend la forme P.P.C.M. $(n, m) = nm$.

3. Algorithme de division dans $\mathbb{Z}$.— *Soient donnés a, $b \in \mathbb{Z}$, avec $b > 0$. Il existe toujours q, $r \in \mathbb{Z}$ tels que*

$$a = bq + r, \quad 0 \leqslant r < b$$

(*si l'on ne pose que $b \neq 0$, c'est l'inégalité $0 \leqslant r < |b|$ qui sera vérifiée*).

En effet, l'ensemble $S = \{a - bs \mid s \in \mathbb{Z},\ a - bs \geqslant 0\}$ est manifestement non vide (par exemple, $a - b(-a^2) > 0$). Il possède donc le plus petit élément; désignons-le par $r = a - bq$. Par hypothèse, on a $r \geqslant 0$. En supposant que $r \geqslant b$, nous aurions obtenu un élément $r - b = a - b(q + 1) \in S$ plus petit que r. Cette contradiction n'est éliminée que pour $r < b$. ■

Le raisonnement peu compliqué que nous venons de développer fournit aussi un procédé (un *algorithme*) permettant de calculer le *quotient* b et le *reste* r en un nombre fini de pas. L'algorithme de division dans $\mathbb{Z}$ est utilisé pour donner une autre définition du P.G.C.D. et donc du P.P.C.M. si l'on tient compte de la relation (2).

A savoir, posons

$$J = \{nu + mv \mid u,\ v \in \mathbb{Z}\}, \tag{3}$$

où n et m sont des entiers donnés, simultanément non nuls.

Choisissons dans J le plus petit élément positif $d = nu_0 + mv_0$ et utilisons l'algorithme de division. Il vient $n = dq + r$, $0 \leqslant r < d$. Vu le choix de d, la relation

$$r = n - dq = n - (nu_0 + mv_0)\, q = n(1 - u_0q) + m(-v_0q) \in J$$

entraîne l'égalité $r = 0$. Par conséquent, $d \mid n$. En opérant de façon identique, on démontre que $d \mid m$. Soit maintenant d' un diviseur quelconque des nombres n et m. Alors,

$$d' \mid n,\ d' \mid m \Rightarrow d' \mid nu_0,\ d' \mid mv_0 \Rightarrow d' \mid (nu_0 + mv_0) \Rightarrow d' \mid d.$$

Ainsi, d possède toutes les propriétés d'un plus grand commun diviseur et donc $d =$ P.G.C.D.(n, m). Nous pouvons ainsi énoncer l'assertion suivante:

Le plus grand commun diviseur de deux entiers n, m, simultanément non nuls, s'écrit toujours sous la forme

$$\text{P.G.C.D.}(n, m) = nu + mv; \quad u,\ v \in \mathbb{Z}. \tag{4}$$

En particulier, les entiers n, m sont premiers entre eux si, et seulement si,

$$nu + mv = 1 \tag{4'}$$

pour certains u, $v \in \mathbb{Z}$. ■

On a vérifié que si les nombres n et m sont premiers entre eux, on a la relation (4′). Réciproquement, si n et m sont tels que la relation (4′) est vérifiée, on a

$$d \mid n,\ d \mid m \Rightarrow d \mid nu,\ d \mid mv \Rightarrow d \mid (nu + mv) \Rightarrow d \mid 1 \Rightarrow d = \pm 1.$$

La démonstration des relations (4) et (4′) est assez constructive. Il convient de prendre un élément positif quelconque de l'ensemble J (voir (3)) et de le réduire ensuite, à l'aide de l'algorithme de division, jusqu'à ce que l'on obtienne le plus petit élément qui sera justement le plus grand commun diviseur.

EXERCICES

1. Tout nombre premier est de la forme $4k + 1$ ou $4k - 1$. En utilisant la propriété de l'ensemble S du n° 1 d'être stable pour la multiplication démontrer que l'ensemble des nombres premiers de la forme $4k - 1$ est infini. (*Indication*. Pour tout entier naturel n, le nombre $4n! - 1$ admet au moins un diviseur premier p de la forme $4k - 1$, tel que $p > n$.)

2. Démontrer qu'il existe infinité de nombres premiers de la forme $4k + 1$, en s'appuyant sur la proposition non triviale suivante (voir chap. 9, § 2, n° 1). Si $n, m \in \mathbb{Z}$, P.G.C.D. $(n, m) = 1$ et, si p est un nombre premier qui divise $n^2 + m^2$, on a $p = 4k + 1$. (*Indication*. Poser $n = 2$ et $m = p_1p_2 \ldots$ $\ldots p_s$, où $p_1, \ldots, p_s$ sont des nombres premiers distincts de la forme $p_i = 4k_i + 1$. Alors, tout diviseur premier p du nombre impair $n^2 + m^2$ est de la forme $4k + 1$, p n'appartenant pas à l'ensemble $\{p_1, p_2, \ldots, p_s\}$.)

3. Si un entier naturel n est divisible exactement par r nombres premiers différents $p_1, \ldots, p_r$, il existe

$$\varphi(n) = n\left(1 - \frac{1}{p_1}\right) \ldots \left(1 - \frac{1}{p_r}\right)$$

nombres inférieurs à n et premiers avec n. La fonction $\varphi: \mathbb{N} \to \mathbb{N}$ s'appelle fonction d'Euler. Vérifier la validité de la formule donnant les valeurs de $\varphi(n)$ pour $n \leqslant 25$ et pour $n = p^m$ (voir aussi chap. 9, § 1, n° 4).

4. En utilisant la formule binomiale, démontrer par récurrence sur n que, si p est un nombre premier, $n^p - n$ est divisible par p pour tout $n \in \mathbb{Z}$. (*Indication*. En cas d'échec, lire le § 4 du chapitre 4, où est donnée une démonstration s'appuyant sur de « hautes » matières.)

CHAPITRE 2

ESPACES VECTORIELS $\mathbb{R}^n$. MATRICES

Les matrices rectangulaires que nous avons introduites au chapitre 1, § 3, se rencontrent si fréquemment qu'elles ont fait naître une branche autonome des mathématiques appelée *théorie des matrices*. Cette théorie a été établie vers le milieu du siècle passé, mais c'est seulement avec le développement de l'algèbre linéaire qu'elle a acquis plus tard sa plénitude et son élégance. Jusqu'à présent, la théorie des matrices demeure un outil d'étude très important et bien adapté tant aux problèmes pratiques qu'aux constructions abstraites des mathématiques modernes. Dans ce qui suit, nous n'exposerons que les résultats les plus simples de la théorie des matrices.

Le titre du présent chapitre peut faire naître l'illusion que nous allons mettre sur le dos de la géométrie la description des êtres purement algébriques. Or, il ne s'agit en réalité que d'une expression commode et économique des propriétés des matrices et des solutions des systèmes linéaires en un langage emprunté à la géométrie. Les notions d'espace, de vecteur, de dépendance linéaire, de rang d'un système, etc., qui sont des notions communément adoptées, sont développées autant que cela est nécessaire pour nos buts immédiats. Quant à l'intuition géométrique, on lui attribue un rôle plus important dans d'autres cours.

D'ailleurs, nous aurons aussi besoin des espaces vectoriels pour pouvoir considérer les applications linéaires dont les matrices sont des compagnons naturels. C'est justement la composition des applications (voir chap. 1, § 4, n° 2) qui conduit par la voie la plus naturelle à la notion de produit de matrices.

§ 1. Espaces vectoriels $\mathbb{R}^n$

1. Motivation.— En étudiant les systèmes d'équations linéaires nous avons eu à considérer des suites de longueur n dont le sens était différent suivant le cas. C'étaient des lignes $(a_{i1}, a_{i2}, \ldots, a_{in})$, $1 \leqslant i \leqslant m$, d'une matrice $A = (a_{ij})$ à m lignes et n colonnes, et des

solutions $(x_1^\circ, x_2^\circ, \ldots, x_n^\circ)$ du système linéaire de matrice A. La réduction d'un système ou d'une matrice à la forme quasi triangulaire, que nous avons décrite au chapitre 1, § 3, comportait, en plus de la transformation élémentaire de type (I), encore deux opérations importantes: la multiplication de la ligne par un nombre et l'addition de deux lignes. On peut effectuer les mêmes opérations sur les solutions d'un système linéaire homogène. En effet, si $(x_1', x_2', \ldots \ldots, x_n')$ et $(x_1'', x_2'', \ldots, x_n'')$ sont deux solutions du système

$$a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = 0, \qquad i = 1, 2, \ldots, m,$$

et α, β sont deux nombres réels quelconques, alors

$$(\alpha x_1' + \beta x_1'', \alpha x_2' + \beta x_2'', \ldots, \alpha x_n' + \beta x_n'')$$

sera aussi solution du système:

$$\begin{aligned} a_{i1}(\alpha x_1' + \beta x_1'') &+ \\ &+ a_{i2}(\alpha x_2' + \beta x_2'') + \ldots + a_{in}(\alpha x_n' + \beta x_n'') = \\ &= \alpha(a_{i1}x_1' + a_{i2}x_2' + \ldots + a_{in}x_n') + \\ &\qquad + \beta(a_{i1}x_1'' + a_{i2}x_2'' + \ldots + a_{in}x_n'') = 0. \end{aligned}$$

D'autre part, toute suite de longueur n, quoi qu'elle exprime, est un élément d'un ensemble « universel » $\mathbb{R}^n$, c'est-à-dire de la puissance cartésienne n-ième de l'ensemble $\mathbb{R}$ des nombres réels. On a donc intérêt à étudier un être général dont les propriétés puissent être étendues aux matrices et aux solutions des systèmes homogènes.

2. Définitions fondamentales.— Soit n un entier naturel fixe quelconque. On appelle *espace vectoriel* $\mathbb{R}^n$ *de dimension* n *sur* $\mathbb{R}$ l'ensemble $\mathbb{R}^n$ (ses éléments sont appelés *vecteurs lignes* ou tout simplement *vecteurs*) muni des opérations d'addition des vecteurs et de multiplication par un *scalaire*, c'est-à-dire par un nombre réel. Les scalaires sont désignés par des lettres minuscules de l'alphabet latin ou grec, et les vecteurs par des lettres majuscules de l'alphabet latin, de même que les matrices. Au fond, un vecteur $X = (x_1, x_2, \ldots \ldots, x_n)$ peut être considéré comme une matrice de type $(1, n)$. Soient $Y = (y_1, y_2, \ldots, y_n)$ un autre vecteur et λ un scalaire. Par définition, on a

$$X + Y = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n),$$
$$\lambda X = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n).$$

Le vecteur nul $(0, 0, \ldots, 0)$ sera désigné dans la suite par 0. On convient d'identifier $\mathbb{R}^1$ à $\mathbb{R}$.

Les propriétés des opérations sur les nombres réels, connues sûrement du lecteur, sont étendues à $\mathbb{R}^n$. Leur énumération, bien qu'ennuyeuse, donne une idée exacte de ce qu'on doit entendre par espace vectoriel abstrait qui sera étudié plus tard dans le cours d'algèbre linéaire et de géométrie:

EV_1: *$X + Y = Y + X$ pour tous les vecteurs X, $Y \in \mathbb{R}^n$ (commutativité de l'addition)*;

EV_2: *$(X + Y) + Z = X + (Y + Z)$, quels que soient $X, Y, Z \in \mathbb{R}^n$ (associativité de l'addition)*;

EV_3: *il existe un vecteur (nul) 0 tel que l'on ait $X + 0 = X$ pour tout $X \in \mathbb{R}^n$*;

EV_4: *tout $X \in \mathbb{R}^n$ admet un vecteur opposé $-X$ tel que $X + (-X) = 0$*;

EV_5: *$1X = X$ pour tout $X \in \mathbb{R}^n$*;

EV_6: *$(\alpha\beta) X = \alpha (\beta X)$ quels que soient α, $\beta \in \mathbb{R}$, et $X \in \mathbb{R}^n$*;

EV_7: *$(\alpha + \beta) X = \alpha X + \beta X$ (distributivité par rapport à l'addition des scalaires)*;

EV_8: *$\alpha (X + Y) = \alpha X + \alpha Y$ (distributivité par rapport à l'addition des vecteurs)*.

Nous admettrons, sans les justifier, l'unicité des vecteurs 0 et $-X$ dont il s'agit dans EV_3 et EV_4, ainsi que les autres conséquences simples qui découlent des propriétés (ou des axiomes, s'il s'agit d'un espace vectoriel quelconque) indiquées ci-dessus, parce qu'elles sont suffisamment évidentes.

Nous avons appelé $\mathbb{R}^n$ espace de dimension n, mais la notion de *dimension*, elle, ne prendra un sens déterminé qu'à la fin de ce paragraphe après une petite préparation. L'origine du terme « espace vectoriel » est expliquée dans le cours de géométrie analytique, où l'on établit une correspondance biunivoque entre les points (vecteurs) du plan cartésien et leurs coordonnées (x, y). Ce sont justement les opérations sur les vecteurs lignes de $\mathbb{R}^2$ qui correspondent à l'addition des vecteurs (règle du parallélogramme) et à leur multiplication par des nombres.

Outre l'espace vectoriel des vecteurs lignes $(x_1, x_2, \ldots, x_n)$ de longueur n, on considère également un espace vectoriel des *vecteurs colonnes* de hauteur n:

$$\left\| \begin{matrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{matrix} \right\| = [x_1, x_2, \ldots, x_n]$$

(pour les désignations voir chap. 1, § 3). On comprend que la différence entre ces deux espaces est purement conventionnelle, mais nous verrons bientôt qu'il est utile d'avoir les deux variantes de l'espace vectoriel. Le contexte permet généralement de comprendre de quels vecteurs, colonnes ou lignes il s'agit, de sorte qu'on n'introduit pas de symboles spéciaux pour leur désignation.

Soit V un sous-ensemble non vide de $\mathbb{R}^n$. On appelle V *sous-espace vectoriel* *) *de* $\mathbb{R}^n$, si

$$X, Y \in V \Rightarrow \alpha X + \beta Y \in V \qquad (1)$$

---

*) Pour l'instant cette définition ne paraît pas assez satisfaisante, mais à la fin du paragraphe nous dirons quelques mots pour la justifier.

pour tous les α, $\beta \in \mathbb{R}$. En particulier, le vecteur nul est toujours contenu dans V. L'ensemble de tous les vecteurs lignes $(x_1, \ldots, x_{n-1}, 0)$ de composante $x_n = 0$ est un sous-espace; on convient de l'identifier à $\mathbb{R}^{n-1}$. Ainsi, nous avons une suite d'inclusions dites canoniques

$$0 \subset \mathbb{R} \subset \mathbb{R}^2 \subset \ldots \subset \mathbb{R}^{n-1} \subset \mathbb{R}^n.$$

Les solutions de l'équation homogène $x_1 + x_2 + \ldots + x_n = 0$ forment un sous-espace de $\mathbb{R}^n$, $n > 1$, qui diffère de l'espace nul et de l'espace $\mathbb{R}^n$ tout entier. Nous indiquerons plus loin d'autres exemples.

3. Combinaisons linéaires. Enveloppe linéaire.— Etant donné $X_1, X_2, \ldots, X_k$ des vecteurs de l'espace vectoriel $\mathbb{R}^n$ et $\alpha_1, \alpha_2, \ldots, \alpha_k$ des scalaires, le vecteur $X = \alpha_1 X_1 + \alpha_2 X_2 + \ldots + \alpha_k X_k$ s'appelle *combinaison linéaire* des vecteurs X_i à coefficients α_i. Par exemple, on a $(2, 3, 5, 5) - 3(1, 1, 1, 1) + 2(1, 0, -1, -1) = (1, 0, 0, 0)$. Soit $Y = \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k$ une autre combinaison linéaire des mêmes vecteurs X_i à coefficients β_i, et $\alpha, \beta \in \mathbb{R}$. Alors

$$\begin{aligned}\alpha X + \beta Y = \alpha(\alpha_1 X_1 + \alpha_2 X_2 + \ldots + \alpha_k X_k) + \\ + \beta(\beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k) = \\ = (\alpha\alpha_1 + \beta\beta_1) X_1 + (\alpha\alpha_2 + \beta\beta_2) X_2 + \ldots + (\alpha\alpha_k + \beta\beta_k) X_k\end{aligned}$$

représente de nouveau une combinaison linéaire des vecteurs X_i à coefficients $\alpha\alpha_i + \beta\beta_i$. On voit que l'*ensemble de toutes les combinaisons linéaires d'un système de vecteurs donnés $X_1, X_2, \ldots, X_k$ est un sous-espace vectoriel de $\mathbb{R}^n$*. On le désigne généralement par le symbole $\langle X_1, X_2, \ldots, X_k \rangle$ et on l'appelle *enveloppe linéaire* du système de vecteurs $X_1, X_2, \ldots, X_k$. On dit encore que le sous-espace $\langle X_1, X_2, \ldots, X_k \rangle$ est *engendré* par les vecteurs $X_1, X_2, \ldots, X_k$.

Il est possible de définir l'enveloppe linéaire de tout sous-ensemble $S \subset \mathbb{R}^n$ en entendant par $\langle S \rangle$ l'ensemble de toutes les combinaisons linéaires d'un système fini de vecteurs de S. Il est clair que si V est un sous-espace de $\mathbb{R}^n$, alors $\langle V \rangle = V$, à savoir: toute combinaison linéaire des vecteurs de V appartient à V. En particulier, $S \subset V \Rightarrow \langle S \rangle \subset V$, c'est-à-dire l'enveloppe linéaire $\langle S \rangle$ peut être définie comme intersection de tous les sous-espaces contenant l'ensemble donné S des vecteurs de $\mathbb{R}^n$:

$$\langle S \rangle = \bigcap_{S \subset V} V. \tag{2}$$

A première vue, il n'est pas immédiat que le second membre de l'égalité (2), représentant l'intersection $\bigcap V$ d'une famille quelconque de sous-espaces, sera un sous-espace. Soient $X, Y \in \bigcap V$, on a alors

X, $Y \in V$, quel que soit le sous-espace V de la famille. Par conséquent, $\alpha X + \beta Y \in V$ pour tous les α, $\beta \in \mathbb{R}$, ce qui donne $\alpha X + \beta Y \in \bigcap V$.

Au contraire, la réunion $U \bigcup V$ des sous-espaces U et V n'est pas en général un sous-espace, comme le montre l'exemple des sous-espaces $U = \{(\lambda, 0) \mid \lambda \in \mathbb{R}\}$ et $V = \{(0, \lambda) \mid \lambda \in \mathbb{R}\}$ de $\mathbb{R}^2$.

L'enveloppe linéaire $\langle U \bigcup V \rangle$ s'appelle *somme* des sous-espaces U et V:

$$U + V = \langle U \bigcup V \rangle = \{u + v \mid u \in U, v \in V\}.$$

Si $U \bigcap V = 0$, on dit que la somme $U + V$ est *directe* et on la note $U \oplus V$. Soient $V = V_1 \oplus V_2$ et $X = X_1 + X_2 = X_1' + X_2'$ deux expressions du vecteur $X \in V$ sous forme de combinaisons linéaires des vecteurs X_1, $X_1' \in V_1$ et X_2, $X_2' \in V_2$. On a alors $X_1 - X_1' = X_2' - X_2 \in V_1 \bigcap V_2$. Puisque $V_1 \bigcap V_2 = 0$, il vient $X_1 = X_1'$, $X_2 = X_2'$. Réciproquement, si l'écriture $X = X_1 + X_2$, $X_i \in V_i$, $i = 1, 2$, est unique pour tout vecteur $X \in V$, la somme $V = V_1 + V_2$ est directe (le lecteur pourra le vérifier à titre d'exercice). En général, la somme V des sous-espaces $V_1, \ldots, V_k \subset \mathbb{R}^n$ est dite directe, $V = V_1 \oplus \ldots \oplus V_k$, si tout vecteur $X \in V$ s'écrit de manière unique sous la forme $X = X_1 + \ldots + X_k$, avec $X_i \in V_i$.

EXEMPLE 1. — Considérons dans $\mathbb{R}^n$ deux ensembles

$$U_m = \{(\lambda_1, \ldots, \lambda_m, 0, \ldots, 0) \mid \lambda_i \in \mathbb{R}\}$$

et

$$V_m = \{(0, \ldots, 0, \lambda_{m+1}, \ldots, \lambda_n) \mid \lambda_i \in \mathbb{R}\},$$

$0 < m < n$. Il est immédiat de vérifier que U_m, V_m sont des sous-espaces de $\mathbb{R}^n$ et que $U_m + V_m = \mathbb{R}^n$ et $U_m \bigcap V_m = 0$. Par suite, $\mathbb{R}^n = U_m \oplus V_m$.

EXEMPLE 2. — Considérons dans $\mathbb{R}^n$ des vecteurs dits *vecteurs lignes unitaires*

$$E_1 = (1, 0, \ldots, 0), \quad E_2 = (0, 1, \ldots, 0), \ldots, E_n = (0, 0, \ldots, 1). \quad (3)$$

Tout vecteur $X = (x_1, x_2, \ldots, x_n)$ s'écrit d'une manière et d'une seule sous la forme $X = x_1E_1 + x_2E_2 + \ldots + x_nE_n$. Par suite,

$$\mathbb{R}^n = \langle E_1 \rangle \oplus \langle E_2 \rangle \oplus \ldots \oplus \langle E_n \rangle.$$

Les vecteurs colonnes unitaires seront désignés par les symboles

$$E^{(1)} = [1, 0, \ldots, 0], \quad E^{(2)} = [0, 1, \ldots, 0], \ldots, E^{(n)} = [0, 0, \ldots, 1]. \quad (3')$$

4. Dépendance linéaire. — Les vecteurs $X_1, \ldots, X_k$ de l'espace $\mathbb{R}^n$ s'appellent *linéairement dépendants* s'il existe k nombres $\alpha_1, \alpha_2, \ldots, \alpha_k$ simultanément non nuls tels que

$$\alpha_1 X_1 + \alpha_2 X_2 + \ldots + \alpha_k X_k = 0 \quad (4)$$

(le second membre exprime le vecteur nul). Nous dirons aussi que la dépendance linéaire (4) est non triviale. Si $\alpha_1 X_1 + \alpha_2 X_2 + \ldots$

$\ldots + \alpha_k X_k = 0 \Rightarrow \alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$, on dira que les vecteurs $X_1, X_2, \ldots, X_k$ sont *linéairement indépendants* (ou forment un système libre).

L'exemple 2 du n° 3 montre que les vecteurs unitaires $E_1, E_2, \ldots$ $\ldots, E_n$ sont linéairement indépendants. Il est évident qu'un vecteur $X \neq 0$ est toujours linéairement indépendant, car $\lambda X = 0, X \neq 0 \Rightarrow$ $\Rightarrow \lambda = 0$. La propriété du système $X_1, \ldots, X_k$ d'être libre n'est aucunement liée à l'ordre des vecteurs, car les termes $\alpha_i X_i$ de l'égalité (4) peuvent être permutés de façon arbitraire.

THÉORÈME 1. — *Les assertions suivantes sont vraies:*

(i) *un système de vecteurs* $\{X_1, \ldots, X_k\}$, *contenant un sous-système linéairement dépendant, est lui-même linéairement dépendant;*

(ii) *toute partie d'un système de vecteurs linéairement indépendants* $\{X_1, \ldots, X_k\}$ *est linéairement indépendante;*

(iii) *parmi les vecteurs linéairement dépendants* $X_1, \ldots, X_k$ *l'un au moins est une combinaison linéaire des autres;*

(iv) *si l'un des vecteurs* $X_1, \ldots, X_k$ *s'exprime linéairement en fonction des autres, les vecteurs* $X_1, \ldots, X_k$ *sont linéairement dépendants;*

(v) *si les vecteurs* $X_1, \ldots, X_k$ *forment un système libre et les vecteurs* $X_1, \ldots, X_k, X$ *sont linéairement dépendants, alors* X *est une combinaison linéaire des vecteurs* $X_1, \ldots, X_k$;

(vi) *si les vecteurs* $X_1, \ldots, X_k$ *sont linéairement indépendants et le vecteur* X_{k+1} *ne peut pas être exprimé en fonction de ces vecteurs, le système* $X_1, \ldots, X_k, X_{k+1}$ *est linéairement indépendant.*

DÉMONSTRATION. — (i). Supposons par exemple que les s premiers vecteurs $X_1, \ldots, X_s$, $s < k$, soient linéairement dépendants, c'est-à-dire

$$\alpha_1 X_1 + \ldots + \alpha_s X_s = 0,$$

où les α_i ne sont pas tous nuls. Alors, en posant $\alpha_{s+1} = \ldots = \alpha_k =$ $= 0$, on obtient une dépendance linéaire non triviale

$$\alpha_1 X_1 + \ldots + \alpha_s X_s + \alpha_{s+1} X_{s+1} + \ldots + \alpha_k X_k = 0.$$

L'assertion (ii) résulte immédiatement de (i) (raisonnement par l'absurde).

(iii). Soit par exemple $\alpha_k \neq 0$ dans la relation (4). Alors

$$X_k = -\frac{\alpha_1}{\alpha_k} X_1 - \ldots - \frac{\alpha_{k-1}}{\alpha_k} X_{k-1}.$$

(iv). Soit par exemple $X_k = \beta_1 X_1 + \ldots + \beta_{k-1} X_{k-1}$. En posant $\alpha_1 = \beta_1, \ldots, \alpha_{k-1} = \beta_{k-1}, \alpha_k = -1$, on obtient la relation (4) avec le coefficient $\alpha_k \neq 0$.

(v) La relation non triviale

$$\beta_1 X_1 + \ldots + \beta_k X_k + \beta X = 0$$

avec $\beta \neq 0$ donne, par suite de (iii), le résultat voulu. Si pourtant $\beta = 0$, on a $\beta_1 = \ldots = \beta_k = 0$, car, par hypothèse, $X_1, \ldots, X_k$ sont linéairement indépendants.

L'assertion (vi) résulte immédiatement de (v). ■

5. Base. Dimension.— Donnons maintenant une définition importante.

Définition. — *Soit V un sous-espace de $\mathbb{R}^n$. Un système de vecteurs $X_1, \ldots, X_r \in V$ est appelé base de V s'il est linéairement indépendant et son enveloppe linéaire coïncide avec V:*

$$\langle X_1, \ldots, X_r \rangle = V.$$

Il résulte de la définition d'une base et de l'enveloppe linéaire d'un système de vecteurs que tout vecteur $X \in V$ s'écrit d'une manière unique sous la forme $X = \alpha_1 X_1 + \ldots + \alpha_r X_r$. Les coefficients $\alpha_1, \ldots, \alpha_r \in \mathbb{R}$ s'appellent *coordonnées* du vecteur X par rapport à la base $X_1, \ldots, X_r$.

Comme nous l'avons déjà vu, les vecteurs unitaires (3) linéairement indépendants engendrent $\mathbb{R}^n$. Par conséquent, $\{E_1, E_2, \ldots, E_n\}$ est une base de l'espace $\mathbb{R}^n$. Cette base dite *canonique* est loin d'être unique dans $\mathbb{R}^n$. Par exemple, les vecteurs

$$E'_1 = E_1,\ E'_2 = E_1 + E_2,\quad E'_3 = E_1 + E_2 + E_3,\ \ldots$$
$$\ldots,\ E'_n = E_1 + E_2 + \ldots + E_n$$

forment, eux aussi, une base de l'espace $\mathbb{R}^n$ (vérifiez-le de façon correcte). D'autre part, il n'est pas encore clair si tout sous-espace vectoriel de $\mathbb{R}^n$ admet une base et, dans l'affirmative, si le nombre de vecteurs de base est constant. Il s'avère que la réponse à ces deux questions est positive. Nos raisonnements seront fondés sur le lemme suivant:

Lemme. — *Soit V un sous-espace de $\mathbb{R}^n$ ayant pour base $X_1, \ldots, X_r$ et soit $Y_1, Y_2, \ldots, Y_s$ un système de vecteurs de V linéairement indépendants. Alors, $s \leqslant r$.*

Démonstration.. — De même que tous les vecteurs de V, les vecteurs $Y_1, \ldots, Y_s$ sont des combinaisons linéaires des vecteurs de base. Soient

$$Y_1 = a_{11}X_1 + a_{21}X_2 + \ldots + a_{r1}X_r,$$
$$Y_2 = a_{12}X_1 + a_{22}X_2 + \ldots + a_{r2}X_r,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$Y_s = a_{1s}X_1 + a_{2s}X_2 + \ldots + a_{rs}X_r,$$

où a_{ij} sont des scalaires quelconques (ce sont les coordonnées des vecteurs Y_j, définies de façon unique, mais pour l'instant, cela nous

importe peu). Raisonnons par l'absurde. Supposons que $s > r$.

Formons une combinaison linéaire des vecteurs Y_j à coefficients x_j:

$$x_1Y_1 + \ldots + x_sY_s = (a_{11}x_1 + a_{12}x_2 + \ldots + a_{1s}x_s)\, X_1 + \ldots$$
$$\ldots + (a_{r1}x_1 + a_{r2}x_2 + \ldots + a_{rs}x_s)\, X_r$$

et considérons un système de r équations linéaires à s inconnues:

$$\begin{gathered} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1s}x_s = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{r1}x_1 + a_{r2}x_2 + \ldots + a_{rs}x_s = 0. \end{gathered}$$

Puisque, par hypothèse, on a $s > r$, nous pouvons appliquer le corollaire 2 du chapitre 1, § 3, d'après lequel notre système admet une solution non nulle $(x_1^\circ, x_2^\circ, \ldots, x_s^\circ)$. Nous obtenons une dépendance linéaire non triviale

$$x_1^\circ Y_1 + x_2^\circ Y_2 + \ldots + x_s^\circ Y_s = 0,$$

qui est en contradiction avec l'hypothèse du lemme. Par conséquent $s \leqslant r$. ■

THÉORÈME 2. — *Tout sous-espace non nul $V \subset \mathbb{R}^n$ admet une base finie. Toutes les bases de l'espace V sont formées d'un même nombre $r \leqslant n$ de vecteurs (ce nombre s'appelle dimension de l'espace V et se note* $\dim_{\mathbb{R}} V$ *ou tout simplement* $\dim V$).

DÉMONSTRATION. — Soit $V \neq 0$. Supposons qu'on ait trouvé dans V un système de vecteurs linéairement indépendants $X_1, \ldots, X_k$, dont X_1 est un vecteur non nul arbitraire de V. Si l'enveloppe linéaire $\langle X_1, \ldots, X_k \rangle$ ne coïncide pas avec V, on choisit dans V un vecteur X_{k+1}, tel que $X_{k+1} \notin \langle X_1, \ldots, X_k \rangle$. En d'autres termes, X_{k+1} n'est pas une combinaison linéaire des vecteurs $X_1, \ldots, X_k$. D'après le théorème 1 (vi), les vecteurs $X_1, \ldots, X_k, X_{k+1}$ sont linéairement indépendants. On pourrait continuer indéfiniment ce processus d'extension du système de vecteurs linéairement indépendants, mais, tous ses vecteurs X_i étant dans $\mathbb{R}^n = \langle E_1, E_2, \ldots, E_n \rangle$, le système contient au plus n vecteurs vu le lemme que nous venons de démontrer. Cela veut dire que, pour un certain entier naturel $r \leqslant n$, le système des vecteurs linéairement indépendants $X_1, \ldots, X_k, \ldots, X_r \in V$ devient maximal, c'est-à-dire que l'on obtient un système linéairement dépendant $X_1, \ldots, X_r, X$, quel que soit le vecteur $X \neq 0$ de V. Suivant le théorème 1 (v) nous aurons $X \in \langle X_1, \ldots, X_r \rangle$. Par conséquent, $V = \langle X_1, \ldots, X_r \rangle$ et les vecteurs $X_1, \ldots, X_r$ forment une base de V.

Supposons maintenant que $Y_1, \ldots, Y_s$ forment une autre base de V. D'après le lemme, on a $s \leqslant r$. En permutant les systèmes $X_1, \ldots, X_r$ et $Y_1, \ldots, Y_s$, il vient, en vertu du même lemme, l'iné-

galité $r \leqslant s$. Par conséquent, $s = r$, et le théorème se trouve donc démontré. ■

Remarquons, bien que cela ne soit pas si nécessaire, que tous nos raisonnements s'appliquent dans la même mesure aussi bien à l'espace des vecteurs lignes qu'à celui des vecteurs colonnes.

Ainsi, à chaque sous-espace vectoriel V de $\mathbb{R}^n$ est associé un entier positif $r \leqslant n$ que nous avons appelé dimension de $V : r = \dim V$. En particulier, $\dim \mathbb{R}^n = n$. Ce paramètre numérique bien important de l'espace peut se caractériser par divers autres procédés (voir exercices). L'une des variantes de définition de la dimension est basée sur la notion de rang d'un système de vecteurs. A savoir, si $\{X_1, X_2, \ldots\}$ est un système quelconque, peut-être infini, de vecteurs de l'espace vectoriel $\mathbb{R}^n$, la dimension de l'enveloppe linéaire $\langle X_1, X_2, \ldots \rangle$, comme nous le savons, est au plus égale à n. Elle s'appelle *rang du système* $\{X_1, X_2, \ldots\}$:

$$\text{rang}\,\{X_1, X_2, \ldots\} = \dim \langle X_1, X_2, \ldots \rangle.$$

Disons quelques mots pour justifier le terme « espace vectoriel ». Choisissons une base quelconque $X_1, \ldots, X_r$ du sous-espace vectoriel $V \subset \mathbb{R}^n$. On a alors $X = \alpha_1 X_1 + \ldots + \alpha_r X_r$ pour tout $X \in V$, et l'ensemble V se trouve donc en bijection avec l'ensemble de toutes les suites de coordonnées $(\alpha_1, \ldots, \alpha_r)$ de longueur r (ou $[\alpha_1, \ldots, \alpha_r]$ de hauteur r). En outre, cette correspondance transforme la combinaison linéaire des vecteurs en une combinaison linéaire des suites. On voit donc que le choix d'une base quelconque dans V nous permet d'interpréter V comme espace vectoriel $\mathbb{R}^r$ de coordonnées inclus d'une certaine façon dans $\mathbb{R}^n$, $n \geqslant r$.

EXERCICES

1. Soient V, V_1 et V_2 trois sous-espaces de $\mathbb{R}^n$ tels que $V \subset V_1 + V_2$. Est-il toujours vrai que $V = V \cap V_1 + V \cap V_2$? Que peut-on dire de cette relation dans le cas particulier où $V_1 \subset V$?

2. Soit V un sous-espace de $\mathbb{R}^n$. Si $V = U \oplus W$ se décompose en somme directe, le sous-espace W s'appelle *supplémentaire* de U, et U supplémentaire de W dans V. Le supplémentaire de U dans V est-il défini de façon unique? Comparer avec la notion de complémentaire $V \setminus U$ adoptée en théorie des ensembles (voir chapitre 1, § 4).

3. Montrer que les vecteurs $X_1 = (1, 2, 3)$, $X_2 = (3, 2, 1)$ sont linéairement indépendants; considérer l'enveloppe linéaire $V = \langle X_1, X_2 \rangle$; montrer que le vecteur $X = (-5, 2, 9)$ est contenu dans V et déterminer ses coordonnées par rapport à la base X_1, X_2; trouver dans $\mathbb{R}^3$ au moins un supplémentaire de V.

4. Montrer qu'un système de vecteurs $X_1, \ldots, X_n$ de $\mathbb{R}^n$ engendre $\mathbb{R}^n$ si, et seulement si, ces vecteurs sont linéairement indépendants.

5. Montrer que tout système de vecteurs linéairement indépendants $X_1, \ldots, X_k$ du sous-espace $V \subset \mathbb{R}^n$ peut être inclus dans un certain système de base de V.

6. Soient U et V deux sous-espaces de $\mathbb{R}^n$. Démontrer que $\dim (U + V) = \dim U + \dim V$ si $U \cap V = 0$.

7. Déterminer le rang du système de vecteurs $(0, 1, 1)$, $(1, 0, 1)$, $(1, 1, 0)$.

§ 2. Rang d'une matrice

1. Retour aux équations.— Etant donné un espace vectoriel $\mathbb{R}^m$ des vecteurs colonnes de hauteur m, considérons n vecteurs

$$A^{(j)} = [a_{1j}, a_{2j}, \ldots, a_{mj}], \quad j = 1, 2, \ldots, n,$$

et leur enveloppe linéaire $V = \langle A^{(1)}, A^{(2)}, \ldots, A^{(n)} \rangle$. Soit donné encore un vecteur $B = [b_1, b_2, \ldots, b_m]$. On demande si B appartient au sous-espace $V \subset \mathbb{R}^m$ et, dans l'affirmative, comment ses coordonnées $b_1, \ldots, b_m$ (par rapport à la base canonique (3'), § 1) sont exprimées en fonction des coordonnées des vecteurs $A^{(j)}$. Dans le cas où $\dim V = n$, la deuxième partie de la question se rapporte aux coordonnées du vecteur B par rapport à la base $A^{(1)}, \ldots$ $\ldots, A^{(n)}$. Prenons une combinaison linéaire des vecteurs $A^{(j)}$ à coefficients arbitraires x_j et établissons l'équation $x_1A^{(1)} + \ldots +$ $+ x_nA^{(n)} = B$. La présentation de cette équation sous la forme

$$x_1 \begin{Vmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{Vmatrix} + x_2 \begin{Vmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{Vmatrix} + \ldots + x_n \begin{Vmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{Vmatrix} = \begin{Vmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{Vmatrix} \qquad (1)$$

n'est qu'une autre écriture du système de m équations linéaires à n inconnues

$$\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m. \end{array} \qquad (2)$$

C'est justement un tel système que nous avons rencontré pour la première fois au chapitre 1, § 3. Nous y avons introduit également une matrice simple et une matrice complète

$$A = \begin{Vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{Vmatrix}, \quad (A \mid B) = \left\Vert \begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array} \right\Vert \qquad (3)$$

du système linéaire (2). Une première impression est telle que nous sommes revenus au point de départ et que nous avons perdu notre temps sans avoir rien gagné. Or, en réalité, nous avons maintenant à notre disposition toute une série de notions importantes. Il ne reste qu'à apprendre à les bien manier.

Maintenant, il est commode de nous entendre sur les désignations. Dans ce qui suit, nous désignerons souvent la somme $s_1 +$

$+ s_2 + \ldots + s_n$ par $\sum_{i=1}^{n} s_i$. Ici $s_1, \ldots, s_n$ sont des grandeurs de nature quelconque (nombres, vecteurs, etc.) qui satisfont à toutes les lois d'addition des nombres ou des vecteurs. Les règles

$$\sum_{i=1}^{n} ts_i = t \sum_{i=1}^{n} s_i, \quad \sum_{i=1}^{n} (s_i + t_i) = \sum_{i=1}^{n} s_i + \sum_{i=1}^{n} t_i$$

sont suffisamment claires pour qu'il soit nécessaire de les expliquer.

Nous allons également considérer des *sommes doubles*:

$$\sum_{j=1}^{n} \sum_{i=1}^{m} a_{ij} = \sum_{j=1}^{n} \left(\sum_{i=1}^{m} a_{ij}\right) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij}\right) = \sum_{i,j} a_{ij}$$

dans lesquelles l'ordre de sommation (suivant le premier ou le deuxième indice) peut être choisi à notre volonté. Cela se conçoit sans peine si l'on dispose les a_{ij} sous la forme d'une matrice rectangulaire de dimensions m et n: la sommation des éléments de cette matrice peut être commencée aussi bien par les lignes que par les colonnes.

D'autres types possibles de sommation seront expliqués à mesure qu'ils se rencontreront.

2. Rang d'une matrice.— On appelle *espace des vecteurs colonnes* de la matrice A à m lignes et n colonnes (voir (3)) l'espace $V = \langle A^{(1)}, A^{(2)}, \ldots, A^{(n)} \rangle$ que nous avons introduit plus haut et que nous désignerons maintenant par le symbole $V_v(A)$ ou tout simplement par V_v (v = vertical). Sa dimension $\mathrm{r}_v(A) = \dim V_v$ sera appelée *rang* de la matrice A *par rapport aux colonnes*. De façon analogue, on introduit le *rang* de la matrice A *par rapport aux lignes*: $\mathrm{r}_h(A) = \dim V_h$, où $V_h = \langle A_1, A_2, \ldots, A_m \rangle$ est un sous-espace de $\mathbb{R}^n$ engendré par les vecteurs lignes $A_i = (a_{i1}, a_{i2}, \ldots, a_{in})$, $i = 1, 2, \ldots, m$ (h = horizontal). Autrement dit

$$\mathrm{r}_v(A) = \mathrm{rang}\,\{A^{(1)}, A^{(2)}, \ldots, A^{(n)}\},$$
$$\mathrm{r}_h(A) = \mathrm{rang}\,\{A_1, A_2, \ldots, A_m\}$$

sont les rangs des systèmes de vecteurs colonnes et de vecteurs lignes respectivement. En vertu du théorème 2 du § 1, $\mathrm{r}_v(A)$ et $\mathrm{r}_h(A)$ sont définis correctement.

Conformément à la définition donnée au chapitre 1, § 3, on dit que la matrice A' est obtenue à partir de A à l'aide d'une *transformation élémentaire de type* (I) si $A'_s = A_t$, $A'_t = A_s$ pour un couple quelconque d'indices $s \neq t$ et $A'_i = A_i$ pour $i \neq s, t$. Si $A'_i = A_i$ pour tous les $i \neq s$ et $A'_s = A_s + \lambda A_t$, $s \neq t$, $\lambda \in \mathbb{R}$, on dit qu'on a appliqué à A une *transformation élémentaire de type* (II).

Remarquons que les transformations élémentaires de deux types sont inversibles, c'est-à-dire que la matrice A', obtenue à partir de A par une transformation élémentaire, se transforme de nouveau

en matrice A, si on lui applique une seule transformation élémentaire du même type.

LEMME. — *Si la matrice A' est obtenue à partir de la matrice rectangulaire A par emploi d'une suite finie de transformations élémentaires, on a les égalités*

(i) $\mathrm{r}_h(A') = \mathrm{r}_h(A)$;
(ii) $\mathrm{r}_v(A') = \mathrm{r}_v(A)$.

DÉMONSTRATION. — Il suffit de considérer le cas où la matrice A' est obtenue en appliquant à A une seule transformation élémentaire (en abrégé t.é).

(i) Il est évident que $\langle A_1, \ldots, A_s, \ldots, A_t, \ldots, A_m \rangle = \langle A_1, \ldots, A_t, \ldots, A_s, \ldots, A_m \rangle$, donc la t.é. de type (I) ne modifie pas $\mathrm{r}_h(A)$. $A'_s = A_s + \lambda A_t \Rightarrow A_s = A'_s - \lambda A_t$, par conséquent, $\langle A_1, \ldots, A_s + \lambda A_t, \ldots, A_t, \ldots, A_m \rangle = \langle A_1, \ldots, A_s, \ldots, A_t, \ldots, A_m \rangle$, si bien que $\mathrm{r}_h(A)$ ne change pas non plus lors de la t.é. de type (II).

(ii) Soient $A'^{(j)}$, $j = 1, \ldots, n$, les vecteurs colonnes de la matrice A'. Si l'on démontre que

$$\sum_{j=1}^{n} \lambda_j A^{(j)} = 0 \Leftrightarrow \sum_{j=1}^{n} \lambda_j A'^{(j)} = 0,$$

alors, à tout système libre de vecteurs colonnes d'une matrice, y compris au système maximal, il correspond un système de vecteurs colonnes linéairement indépendants, *de mêmes indices*, de l'autre matrice, et l'égalité $\mathrm{r}_v(A') = \mathrm{r}_v(A)$ se trouve vérifiée. Signalons encore que, les transformations élémentaires étant inversibles, il suffit de démontrer l'implication dans un seul sens. Soit, par exemple, $\sum_{j=1}^{n} \lambda_j A^{(j)} = 0$. Alors, en remplaçant dans (1) x_j par λ_j et tous les b_i par 0, nous voyons que $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ est solution du système homogène (en abrégé SH) associé au système linéaire (2). En vertu du théorème 1 du chapitre 1, cette solution sera aussi solution de SH' obtenu à partir de SH à l'aide de t.é. de type (I) ou (II) et ayant pour matrice justement la matrice A'. Puisque le système SH' s'écrit en abrégé sous forme de $\sum_{j=1}^{n} x_j A'^{(j)} = 0$, nous obtenons la relation $\sum_{j=1}^{n} \lambda_j A'^{(j)} = 0$. ■

Le résultat principal du présent paragraphe est énoncé dans le théorème suivant:

THÉORÈME 1. — *Pour toute matrice rectangulaire A à m lignes et n colonnes, on a la relation* $\mathrm{r}_v(A) = \mathrm{r}_h(A)$ (*ce nombre s'appelle simplement rang de la matrice A et se note* rang A).

DÉMONSTRATION. — Appliquant aux lignes A_i un nombre fini de transformations élémentaires, on peut réduire la matrice A à une forme quasi triangulaire (théorème 2 du chap. 1, § 3)

$$\bar{A} = \begin{Vmatrix} \bar{a}_{11} & \ldots & \bar{a}_{1k} & \ldots & \bar{a}_{1l} & \ldots & \bar{a}_{1s} & \ldots & \bar{a}_{1n} \\ 0 & \ldots & \bar{a}_{2k} & \ldots & \bar{a}_{2l} & \ldots & \bar{a}_{2s} & \ldots & \bar{a}_{2n} \\ 0 & \ldots & 0 & \ldots & \bar{a}_{3l} & \ldots & \bar{a}_{3s} & \ldots & \bar{a}_{3n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \ldots & 0 & \ldots & 0 & \ldots & \bar{a}_{rs} & \ldots & \bar{a}_{rn} \\ 0 & \ldots & 0 & \ldots & 0 & \ldots & 0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \ldots & 0 & \ldots & 0 & \ldots & 0 & \ldots & 0 \end{Vmatrix} \qquad (4)$$

avec $\bar{a}_{11}\bar{a}_{2k}\bar{a}_{3l} \ldots \bar{a}_{rs} \neq 0$. D'après le lemme on a

$$\mathrm{r}_v(A) = \mathrm{r}_v(\bar{A}), \quad \mathrm{r}_h(A) = \mathrm{r}_h(\bar{A}),$$

si bien qu'il suffit de démontrer seulement l'égalité $\mathrm{r}_v(\bar{A}) = \mathrm{r}_h(\bar{A})$.

Les vecteurs colonnes des matrices A et $\bar{A}$ d'indices 1, k, l, ..., s, associés aux inconnues principales $x_1, x_k, x_l, \ldots, x_s$ du système linéaire (2), sont appelés *vecteurs colonnes de base*. Cette terminologie est parfaitement justifiée. En supposant qu'il existe une relation

$$\lambda_1\bar{A}^{(1)} + \lambda_k\bar{A}^{(k)} + \lambda_l\bar{A}^{(l)} + \ldots + \lambda_s\bar{A}^{(s)} = 0$$

entre les vecteurs colonnes $\bar{A}^{(1)} = [\bar{a}_{11}, 0, \ldots, 0]$, $\bar{A}^{(k)} = [\bar{a}_{1k}, \bar{a}_{2k}, 0, \ldots, 0]$, $\bar{A}^{(l)} = [\bar{a}_{1l}, \bar{a}_{2l}, \bar{a}_{3l}, 0, \ldots, 0]$, ..., $\bar{A}^{(s)} = [\bar{a}_{1s}, \bar{a}_{2s}, \ldots, \bar{a}_{rs}, 0, \ldots, 0]$ de la matrice (4), nous obtenons successivement $\lambda_s\bar{a}_{rs} = 0, \ldots, \lambda_l\bar{a}_{3l} = 0$, $\lambda_k\bar{a}_{2k} = 0$, $\lambda_1\bar{a}_{11} = 0$. Puisque $\bar{a}_{11}\bar{a}_{2k} \ldots \bar{a}_{rs} \neq 0$, on a $\lambda_1 = \lambda_k = \lambda_l = \ldots = \lambda_s = 0$. Cela signifie que rang $\{\bar{A}^{(1)}, \bar{A}^{(k)}, \bar{A}^{(l)}, \ldots, \bar{A}^{(s)}\} = r$ et $\mathrm{r}_v(\bar{A}) \geqslant r$. Or, l'espace $\bar{V}_v$, engendré par les vecteurs colonnes de la matrice $\bar{A}$, s'identifie à l'espace des vecteurs colonnes de la matrice obtenue à partir de $\bar{A}$ en supprimant les $m - r$ dernières lignes nulles. C'est pourquoi $\mathrm{r}_v(\bar{A}) = \dim V_v \leqslant \dim \mathbb{R}^r = r$. La comparaison de deux inégalités montre que $\mathrm{r}_v(\bar{A}) = r$ (l'inégalité $\mathrm{r}_v(\bar{A}) \leqslant r$ découle aussi du fait évident que tous les vecteurs colonnes de la matrice $\bar{A}$ sont des combinaisons linéaires des vecteurs colonnes de base; nous laissons au lecteur le soin de le vérifier à titre d'exercice).

D'autre part, tous les vecteurs lignes non nuls de la matrice $\bar{A}$ sont linéairement indépendants; toute relation hypothétique

$$\lambda_1\bar{A}_1 + \lambda_2\bar{A}_2 + \ldots + \lambda_r\bar{A}_r = 0, \quad \lambda_i \in \mathbb{R},$$

donne successivement, de même que dans le cas des vecteurs colonnes, $\lambda_1\bar{a}_{11} = 0$, $\lambda_2\bar{a}_{2k} = 0$, . . ., $\lambda_r\bar{a}_{rs} = 0$, d'où $\lambda_1 = \lambda_2 = \ldots$ $\ldots = \lambda_r = 0$. Par conséquent, $\mathrm{r}_h(\bar{A}) = r = \mathrm{r}_v(\bar{A})$. ■

3. Critère de compatibilité.— La forme quasi triangulaire de la matrice A, qui permet de répondre à certaines questions relatives aux systèmes linéaires (voir chap. 1, § 3), contient des éléments de l'arbitraire, dus par exemple au choix des colonnes de base ou, ce qui revient au même, au choix des inconnues principales du système (2). En même temps, le théorème 1 et sa démonstration entraînent l'assertion suivante:

COROLLAIRE. — *Le nombre d'inconnues principales du système linéaire* (2) *ne dépend pas du procédé utilisé pour le réduire à la forme quasi triangulaire et est égal au* rang A, *où* A *est la matrice du système.*

En effet, nous avons vu que le nombre d'inconnues principales est égal à celui de lignes non nulles de la matrice $\bar{A}$ (voir (4)) et coïncide avec le rang de la matrice A. Quant au rang, nous l'avons défini d'une façon parfaitement invariante. Cela veut dire que le rang d'une matrice représente sa caractéristique interne qui ne dépend pas des circonstances accessoires. ■

Au chapitre 3 nous obtiendrons un moyen efficace permettant de calculer le rang d'une matrice A sans qu'il soit nécessaire de la réduire à la forme quasi triangulaire. Cela va conférer évidemment plus d'importance aux assertions fondées sur la notion de rang. Enonçons, à titre d'un exemple bien simple mais utile, le critère de compatibilité d'un système linéaire dont il s'agissait encore au chapitre 1.

THÉORÈME 2 (de Kronecker-Capelli). — *Un système d'équations linéaires* (2) *est compatible si, et seulement si, le rang de sa matrice coïncide avec le rang de la matrice complète* (voir (4)).

DÉMONSTRATION. — La compatibilité du système linéaire (2), écrit sous la forme (1), peut être interprétée (on a commencé par là le présent paragraphe) comme une question relative à la représentation du vecteur colonne B des termes constants sous la forme d'une combinaison linéaire des vecteurs colonnes $A^{(j)}$ de la matrice A. Si une telle représentation est possible (c'est-à-dire, si le système (2) est compatible), on a $B \in \langle A^{(1)}, \ldots, A^{(n)} \rangle$ et rang $\{A^{(1)}, \ldots$ $\ldots, A^{(n)}\} =$ rang $\{A^{(1)}, \ldots, A^{(n)}, B\}$, d'où rang $A = \mathrm{r}_v(A) =$ $= \mathrm{r}_v((A \mid B)) =$ rang $(A \mid B)$ (voir l'énoncé du théorème 1).

Réciproquement, si les rangs des matrices A et $(A \mid B)$ coïncident et $\{A^{(j_1)}, \ldots, A^{(j_r)}\}$ est un système libre maximal quelconque que forment les vecteurs colonnes de la matrice A, le système complet $\{A^{(j_1)}, \ldots, A^{(j_r)}, B\}$ sera linéairement dépendant, ce qui

signifie, en vertu du théorème 1 (v) du § 1, que B est une combinaison linéaire des vecteurs colonnes de base, et à plus forte raison, de tous les vecteurs colonnes $A^{(j)}$. Par conséquent, le système (2) est compatible. ■

EXERCICES

1. Démontrer le théorème 1 sans réduire la matrice $A = (a_{ij})$ de type (m, n) à la forme quasi triangulaire. (I n d i c a t i o n. Soient $\dim V_h(A) = r$, $\dim V_v(A) = s$. Choisir r vecteurs lignes de base; sans restreindre la généralité, on peut considérer les r premiers vecteurs lignes $A_1, A_2, \ldots, A_r$. Considérer la matrice $\widetilde{A} = [A_1, A_2, \ldots, A_r]$ de type (r, n) formée de r premières lignes de la matrice A. Choisir dans $\widetilde{A}$ t vecteurs colonnes de base, $t = \dim V_v(\widetilde{A})$; soit $\widetilde{A}^{(1)}, \ldots, \widetilde{A}^{(t)}$. Puisque $V_v(\widetilde{A}) \subset \mathbb{R}^r$, on a $t \leqslant r$. Trouver pour chaque vecteur colonne $A^{(k)}$, $k > t$, des scalaires $\lambda_1, \ldots, \lambda_t \in \mathbb{R}$, tels que $A^{(k)} = \lambda_1 A^{(1)} + \ldots + \lambda_t A^{(t)}$, c'est-à-dire $a_{ik} = \sum_{p=1}^{t} \lambda_p a_{ip}$, $1 \leqslant i \leqslant m$. Pour $i \leqslant r$, il en sera sûrement ainsi, car les vecteurs colonnes raccourcis vérifient la relation $\widetilde{A}^{(k)} = \lambda_1 \widetilde{A}^{(1)} + \ldots + \lambda_t \widetilde{A}^{(t)}$. Pour $i > r$, utiliser l'expression $A_i = \mu_1 A_1 + \ldots + \mu_r A_r$ du vecteur ligne d'indice i en fonction des r premiers vecteurs lignes. On en tire $a_{ik} = \sum_{l=1}^{r} \mu_l a_{lk} = \sum_{l=1}^{r} \mu_l \sum_{p=1}^{t} \lambda_p a_{lp} = \sum_{p=1}^{t} \lambda_p \sum_{l=1}^{r} \mu_l a_{lp} = \sum_{p=1}^{t} \lambda_p a_{ip}$. La dépendance linéaire des vecteurs colonnes ainsi établie montre que $s \leqslant t$. Puisque $t \leqslant r$, on a $s \leqslant r$. Considérer ensuite une matrice dite *transposée*

$$
{}^tA = \begin{Vmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \cdot & \cdot & \cdot & \cdot \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{Vmatrix}
$$

ayant n lignes et m colonnes. On a les égalités $\mathrm{r}_h({}^tA) = \mathrm{r}_v(A)$, $\mathrm{r}_v({}^tA) = \mathrm{r}_h(A)$. Par suite de ce qui a été démontré, $r \leqslant s$. Donc, $r = s$.)

2. De même que dans le cas des lignes, la permutation des colonnes de numéros s et t d'une matrice A est appelée transformation élémentaire (t.é.) de type (I), alors que l'addition à la s-ième colonne de la t-ième colonne multipliée par un scalaire λ, est appelée transformation élémentaire de type (II). Indiquer la forme quasi triangulaire de la matrice A suivant les colonnes. En appliquant des transformations élémentaires à ses colonnes, réduire la matrice $\overline{A}$ (voir (4)) à la forme

$$
\widetilde{A} = \begin{Vmatrix} \tilde{a}_{11} & & & & & & & 0 \\ & \tilde{a}_{22} & & & & & & \\ & & \ddots & & & & & \\ & & & \tilde{a}_{rr} & & & & \\ & & & & 0 & & & \\ & & & & & \ddots & & \\ 0 & & & & & & & 0 \end{Vmatrix},
$$

où $\tilde{a}_{11}=\bar{a}_{11}$, $\tilde{a}_{22}=\bar{a}_{2k}$, $\tilde{a}_{33}=\bar{a}_{3l}$, ..., $\tilde{a}_{rr}=\bar{a}_{rs}$; $\prod_{i=1}^{r} \tilde{a}_{ii} \neq 0$.

3. Montrer que pour $a_0 \neq 0$ la matrice carrée

$$A=\begin{Vmatrix} 0 & 0 & \dots & 0 & 0 & a_0 \\ 1 & 0 & \dots & 0 & 0 & a_1 \\ 0 & 1 & \dots & 0 & 0 & a_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & 1 & 0 & a_{n-1} \\ 0 & 0 & \dots & 0 & 1 & a_n \end{Vmatrix}$$

est de rang n.

4. Exprimer la condition d'égalité des rangs de deux matrices

$$A=\begin{Vmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_n \\ \beta_1 & \beta_2 & \dots & \beta_n \end{Vmatrix}, \qquad B=\begin{Vmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_n \\ \beta_1 & \beta_2 & \dots & \beta_n \\ \gamma_1 & \gamma_2 & \dots & \gamma_n \end{Vmatrix}$$

par une propriété géométrique de l'ensemble de n droites du plan.

§ 3. Applications linéaires. Opérations sur les matrices

1. Matrices et applications.— Soient $\mathbb{R}^n$ et $\mathbb{R}^m$ deux espaces vectoriels de vecteurs colonnes de hauteurs respectives n et m. Soit ensuite $A = (a_{ij})$ une matrice à m lignes et n colonnes. Définissons une application φ_A: $\mathbb{R}^n \to \mathbb{R}^m$ en posant pour tout $X = [x_1, x_2, \dots, x_n] \in \mathbb{R}^n$

$$\varphi_A(X) = x_1A^{(1)} + x_2A^{(2)} + \dots + x_nA^{(n)}, \tag{1}$$

où $A^{(1)}, \dots, A^{(n)}$ sont les vecteurs colonnes de la matrice A (comparer à (1) du § 2). Ces derniers étant de hauteur m, le second membre de (1) est un vecteur colonne $Y = [y_1, y_2, \dots, y_m] \in \mathbb{R}^m$. Ecrivons l'expression (1) sous une forme plus détaillée

$$y_i = \sum_{j=1}^{n} a_{ij}x_j, \quad i=1, 2, \dots, m. \tag{1'}$$

Si $X = X' + X'' = [x_1' + x_1'', x_2' + x_2'', \dots, x_n' + x_n'']$, on a

$$\varphi_A(X'+X'') = \sum_{i=1}^{n} (x_i' + x_i'') A^{(i)} = \sum_{i=1}^{n} x_i' (A^{(i)}) + \sum_{i=1}^{n} x_i'' A^{(i)} = \\ = \varphi_A(X') + \varphi_A(X'').$$

De façon analogue

$$\varphi_A(\lambda X) = \sum_{i=1}^{n} \lambda x_i A^{(i)} = \lambda \sum_{i=1}^{n} x_i A^{(i)} = \lambda\varphi_A(X), \quad \lambda \in \mathbb{R}.$$

Réciproquement, supposons que $\varphi: \mathbb{R}^n \to \mathbb{R}^m$ est une application d'un ensemble dans un autre ensemble (voir chap. 1, § 5) qui vérifie deux propriétés suivantes:

(i) $\varphi(X' + X'') = \varphi(X') + \varphi(X'')$ *pour tous les* $X', X'' \in \mathbb{R}^n$;

(ii) $\varphi(\lambda X) = \lambda\varphi(X)$ *pour tous les* $X \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$.

Désignons par $E_n^{(1)}, \ldots, E_n^{(n)}$ et $E_m^{(1)}, \ldots, E_m^{(m)}$ les vecteurs colonnes des bases canoniques respectives des espaces $\mathbb{R}^n$ et $\mathbb{R}^m$ (voir § 1, n° 3) et utilisons les propriétés (i), (ii) en les appliquant à un vecteur colonne arbitraire $X = [x_1, x_2, \ldots, x_n] = \sum_{j=1}^{n} x_j E_n^{(j)} \in \mathbb{R}^n$. Il vient

$$\varphi(X) = \varphi\left(\sum_{j=1}^{n} x_j E_n^{(j)}\right) = \sum_{j=1}^{n} x_j \varphi(E_n^{(j)}). \tag{2}$$

La relation (2) montre que l'application φ est bien définie par ses valeurs sur les vecteurs colonnes de base. On pose

$$\varphi(E_n^{(j)}) = \sum_{i=1}^{m} a_{ij} E_m^{(i)} = [a_{1j}, a_{2j}, \ldots, a_{mj}] = A^{(j)} \in \mathbb{R}^m \tag{3}$$

et l'on constate que la donnée de φ est équivalente à celle de la matrice $A = (a_{ij})$ à m lignes et n colonnes $A^{(1)}, \ldots, A^{(n)}$, et qu'en réalité les relations (1) et (2) coïncident. On peut donc poser $\varphi = \varphi_A$.

Définition. — *L'application* $\varphi = \varphi_A: \mathbb{R}^n \to \mathbb{R}^m$ *s'appelle application linéaire de* $\mathbb{R}^n$ *dans* $\mathbb{R}^m$, *si elle vérifie les propriétés* (i), (ii). *En particulier, on dit que* φ *est une transformation linéaire de* $\mathbb{R}^n$, *si* $n = m$. *La matrice* A *s'appelle matrice de l'application linéaire* φ_A.

Soient φ_A, $\varphi_{A'}$ deux applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}^m$ de matrices $A = (a_{ij})$ et $A' = (a'_{ij})$. On a $\varphi_A = \varphi_{A'}$ si, et seulement si, $\varphi_A(X) = \varphi_{A'}(X)$ pour tout $X \in \mathbb{R}^n$. En particulier, $A'^{(j)} = \varphi_{A'}(E_n^{(j)}) = \varphi_A(E_n^{(j)}) = A^{(j)}$, $1 \leqslant j \leqslant n$, d'où $a'_{ij} = a_{ij}$ et $A' = A$.

Résumons nos résultats:

Théorème 1. — *Les applications linéaires de* $\mathbb{R}^n$ *dans* $\mathbb{R}^m$ *sont en correspondance biunivoque avec les matrices à m lignes et n colonnes.* ■

Il est à noter qu'il n'y a aucun sens de parler des applications linéaires $S \to T$ des ensembles quelconques S et T. Les conditions (i), (ii) supposent que S et T sont des sous-espaces des espaces vectoriels $\mathbb{R}^n$ et $\mathbb{R}^m$.

Signalons le cas spécial de $m = 1$, où une application linéaire $\varphi: \mathbb{R}^n \to \mathbb{R}$, qu'on appelle généralement *fonction linéaire* de n

variables, se définit par n scalaires $a_1, a_2, \ldots, a_n$:

$$\varphi(X) = \varphi(x_1, x_2, \ldots, x_n) = a_1x_1 + a_2x_2 + \ldots + a_nx_n. \quad (4)$$

Cette terminologie diffère de celle adoptée à l'école secondaire, où (dans le cas d'une seule variable x) on appelle fonction linéaire une fonction $x \mapsto ax + b$.

On définit pour les fonctions linéaires (4), de même que pour les applications linéaires arbitraires $\mathbb{R}^n \to \mathbb{R}^m$ (n et m sont fixes) les opérations d'addition et de multiplication par des scalaires. Soient φ_A, φ_B: $\mathbb{R}^n \to \mathbb{R}^m$ deux applications linéaires. L'application

$$\varphi = \alpha\varphi_A + \beta\varphi_B: \mathbb{R}^n \to \mathbb{R}^m, \quad \alpha, \beta \in \mathbb{R},$$

est définie par ses valeurs

$$\varphi(X) = \alpha\varphi_A(X) + \beta\varphi_B(X).$$

Le second membre de cette égalité est une combinaison linéaire ordinaire des vecteurs colonnes.

Puisque

$$\begin{aligned}\varphi(X' + X'') &= \alpha\varphi_A(X' + X'') + \beta\varphi_B(X' + X'') = \\ &= \alpha\{\varphi_A(X') + \varphi_A(X'')\} + \beta\{\varphi_B(X') + \varphi_B(X'')\} = \\ &= \{\alpha\varphi_A(X') + \beta\varphi_B(X')\} + \{\alpha\varphi_A(X'') + \beta\varphi_B(X'')\} = \\ &= \varphi(X') + \varphi(X'');\end{aligned}$$

$$\begin{aligned}\varphi(\lambda X) &= \alpha\varphi_A(\lambda X) + \beta\varphi_B(\lambda X) = \alpha\lambda\varphi_A(X) + \beta\lambda\varphi_B(X) = \\ &= \lambda\{\alpha\varphi_A(X) + \beta\varphi_B(X)\} = \lambda\varphi(X)\end{aligned}$$

(nous avons utilisé ici implicitement les règles EV_1 à EV_8 du § 1), φ est une application linéaire. En raison du théorème 1, on peut parler de sa matrice C: $\varphi = \varphi_C$. Pour la déterminer, utilisons (3) et écrivons le vecteur colonne d'indice j sous forme de

$$\begin{aligned}[c_{1j}, c_{2j}, \ldots, c_{mj}] &= C^{(j)} = \varphi_C(E_n^{(j)}) = \\ &= \alpha\varphi_A(E_n^{(j)}) + \beta\varphi_B(E_n^{(j)}) = \alpha A^{(j)} + \beta B^{(j)} = \\ &= [\alpha a_{1j} + \beta b_{1j}, \alpha a_{2j} + \beta b_{2j}, \ldots, \alpha a_{mj} + \beta b_{mj}].\end{aligned}$$

Il est naturel de dire que la matrice $C = (c_{ij})$ d'éléments $c_{ij} = \alpha a_{ij} + \beta b_{ij}$ est une combinaison linéaire des matrices A et B à coefficients α et β:

$$\alpha \begin{Vmatrix} a_{11} & \ldots & a_{1n} \\ \cdot & \cdot\cdot\cdot & \cdot \\ a_{m1} & \ldots & a_{mn} \end{Vmatrix} + \beta \begin{Vmatrix} b_{11} & \ldots & b_{1n} \\ \cdot & \cdot\cdot\cdot & \cdot \\ b_{m1} & \ldots & b_{mn} \end{Vmatrix} = \begin{Vmatrix} \alpha a_{11} + \beta b_{11} & \ldots & \alpha a_{1n} + \beta b_{1n} \\ \cdot & \cdot\cdot\cdot & \cdot \\ \alpha a_{m1} + \beta b_{m1} & \ldots & \alpha a_{mn} + \beta b_{mn} \end{Vmatrix}. \quad (5)$$

Ainsi

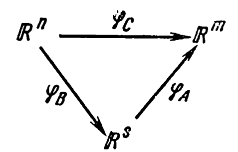
$$\alpha\varphi_A + \beta\varphi_B = \varphi_{\alpha A + \beta B}. \tag{6}$$

Nous profiterons souvent du fait que les combinaisons linéaires des fonctions linéaires sont encore des fonctions linéaires.

Remarquons, avant de clore ce point, que si l'on récrit les propriétés EV_1 à EV_8 établies au § 1 pour les espaces vectoriels, en remplaçant partout les vecteurs lignes X, Y, Z par des matrices de type (m, n), il vient par suite de la relation (5) que l'ensemble des matrices à m lignes et n colonnes forme un espace vectoriel. Si l'on veut, on peut le considérer comme écriture compacte de l'espace vectoriel $\mathbb{R}^{mn}$ des lignes de longueur mn (les lignes sont divisées en tronçons de longueur n, situés l'un au-dessous de l'autre).

2. Produit de matrices.— Les relations (5) et (6) expriment la concordance des opérations d'addition et de multiplication par des scalaires dans les ensembles des matrices à m lignes et n colonnes et des applications $\mathbb{R}^n \to \mathbb{R}^m$. Dans le cas des ensembles arbitraires, il existe aussi une notion importante de composition d'applications (voir chap. 1, § 5, n° 2). Il est logique de s'attendre à ce que la composée (le produit) de deux applications linéaires doive s'exprimer d'une certaine façon en termes de matrices. Voyons comment cela se fait.

Soient $\varphi_B : \mathbb{R}^n \to \mathbb{R}^s$, $\varphi_A : \mathbb{R}^s \to \mathbb{R}^m$ deux applications linéaires et $\varphi_C = \varphi_A \circ \varphi_B$ leur composée :



Il est immédiat de vérifier que $\varphi = \varphi_A \circ \varphi_B$ est une application linéaire. En effet on a :

(i) $\varphi(X' + X'') = \varphi_A(\varphi_B(X' + X'')) =$

$$= \varphi_A(\varphi_B(X') + \varphi_B(X'')) = \varphi_A(\varphi_B(X')) + \varphi_A(\varphi_B(X'')) = \\ = \varphi(X') + \varphi(X'');$$

(ii) $\varphi(\lambda X) = \varphi_A(\varphi_B(\lambda X)) = \varphi_A(\lambda\varphi_B(X)) =$

$$= \lambda\varphi_A(\varphi_B(X)) = \lambda\varphi(X).$$

C'est pourquoi, on associe à φ une matrice C bien définie (théorème 1).

L'action des applications sur les vecteurs colonnes dans la suite

$$[x_1, \ldots, x_n] \overset{\varphi_B}{\longmapsto} [y_1, \ldots, y_s] \overset{\varphi_A}{\longmapsto} [z_1, \ldots, z_m]$$

s'écrit sous forme explicite d'après la formule (1'):

$$z_i = \sum_{k=1}^{s} a_{ik} y_k = \sum_{k=1}^{s} a_{ik} \sum_{j=1}^{n} b_{kj} x_j = \sum_{j=1}^{n} \left(\sum_{k=1}^{s} a_{ik} b_{kj} \right) x_j.$$

D'autre part, on a

$$z_i = \sum_{j=1}^{n} c_{ij} x_j, \quad i = 1, 2, \ldots, m.$$

En comparant les expressions ainsi obtenues et ayant en vue que x_j ($j = 1, 2, \ldots, n$) sont des nombres réels arbitraires, nous obtenons les relations

$$c_{ij} = \sum_{k=1}^{s} a_{ik} b_{kj}, \quad 1 \leqslant i \leqslant m, \quad 1 \leqslant j \leqslant n. \tag{7}$$

On dit que la matrice $C = (c_{ij})$ est obtenue *en multipliant* la matrice A par la matrice B, et l'on écrit:

$$C = AB.$$

Ainsi, on appelle *produit* d'une matrice (a_{ik}) à m lignes et s colonnes par une matrice (b_{kj}) à s lignes et n colonnes, une matrice (c_{ij}) à m lignes et n colonnes, dont les éléments c_{ij} sont définis par la relation (7). Nous avons ainsi démontré le théorème suivant:

THÉORÈME 2. — *Le produit $\varphi_A \varphi_B$ de deux applications linéaires associées aux matrices A et B est une application linéaire associée à la matrice $C = AB$. Autrement dit,*

$$\varphi_A \varphi_B = \varphi_{AB}. \quad \blacksquare \tag{8}$$

La relation (8) complète naturellement la relation (6).

Nous pouvons faire abstraction des applications linéaires et calculer le produit AB de deux matrices quelconques A, B, ayant en vue toutefois que le *symbole AB n'a son sens que dans le cas où le nombre de colonnes de la matrice A est égal au nombre de lignes de la matrice B.* C'est justement à cette condition que fonctionne la règle (7) de « multiplication de la i-ième ligne A_i par la j-ième colonne $B^{(j)}$», d'après laquelle

$$c_{ij} = (a_{i1}, \ldots, a_{is}) [b_{1j}, \ldots, b_{sj}] = A_i B^{(j)}. \tag{9}$$

Le nombre de lignes de la matrice AB est égal au nombre de lignes de la matrice A, alors que le nombre de colonnes est égal à celui de la matrice B. Par conséquent, le produit de matrices carrées de même ordre se trouve toujours défini, mais même dans ce cas on a, en général, $AB \neq BA$ comme le montre l'exemple suivant:

$$\begin{Vmatrix} 1 & 0 \\ 0 & 0 \end{Vmatrix} \begin{Vmatrix} 0 & 0 \\ 1 & 0 \end{Vmatrix} = \begin{Vmatrix} 0 & 0 \\ 0 & 0 \end{Vmatrix} \neq \begin{Vmatrix} 0 & 0 \\ 1 & 0 \end{Vmatrix} = \begin{Vmatrix} 0 & 0 \\ 1 & 0 \end{Vmatrix} \begin{Vmatrix} 1 & 0 \\ 0 & 0 \end{Vmatrix}.$$

Certes, la multiplication des matrices aurait pu être introduite par de nombreux autres procédés (en multipliant, par exemple, ligne par ligne), mais aucun d'eux n'offre tant de possibilités que le procédé décrit ci-dessus. Cela se conçoit aisément, car nous y sommes arrivés d'une façon naturelle en étudiant la composition des applications; or, la notion d'application, elle, est l'une des plus fondamentales en mathématiques.

COROLLAIRE. — *La multiplication matricielle est associative*:

$$A\,(BC) = (AB)\,C.$$

En effet, le produit de matrices correspond au produit d'applications linéaires (le théorème 2 et la relation (8)). Or, d'après le théorème 1 du chapitre 1, § 5, la multiplication des applications est associative. On retrouve le même résultat par calcul, en utilisant directement la relation (7). ■

3. Matrices carrées.— Soit $M_n(\mathbb{R})$ (ou M_n) un ensemble de toutes les matrices carrées (a_{ij}) d'ordre n à coefficients réels a_{ij}.

A l'application identique $e_{\mathbb{R}^n}: \mathbb{R}^n \to \mathbb{R}^n$, qui à tout vecteur colonne $X \in \mathbb{R}^n$ associe ce même vecteur, il correspond évidemment une matrice unité

$$E = \begin{Vmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ 0 & 0 & \dots & 1 \end{Vmatrix}.$$

On peut écrire $E = (\delta_{kj})$, où

$$\delta_{kj} = \begin{cases} 1 & \text{si } k = j, \\ 0 & \text{si } k \neq j \end{cases}$$

est le *symbole de Kronecker*. La règle (7) de multiplication matricielle, où il convient de remplacer b_{kj} par δ_{kj}, montre que les relations suivantes sont vraies:

$$EA = A = AE, \quad \forall A \in M_n(\mathbb{R}). \tag{10}$$

Les relations matricielles (10), obtenues par voie de calcul, découlent, certes, des relations $e\varphi = \varphi = \varphi e$ (voir chap. 1, § 5, n° 2) si l'on utilise le théorème 1 et l'égalité (8), avec $\varphi_A = \varphi$, $\varphi_B = \varphi_E = e$, φ étant une application linéaire quelconque.

Comme nous le savons (voir (5)), les matrices de $M_n(\mathbb{R})$ peuvent être multipliées par des nombres, en entendant par λA, où $A = (a_{ij})$, une matrice (λa_{ij}).

Mais la multiplication par un scalaire (un nombre) se ramène à la multiplication des matrices

$$\lambda A = \operatorname{diag}_n(\lambda)\cdot A = A \operatorname{diag}_n(\lambda), \tag{11}$$

où

$$\operatorname{diag}_n(\lambda) = \lambda E = \begin{Vmatrix} \lambda & 0 & \dots & 0 \\ 0 & \lambda & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & \lambda \end{Vmatrix}$$

est une matrice scalaire que nous avons déjà rencontrée (voir chap. 1, § 4).

L'égalité (11) exprime un fait facile à vérifier, à savoir : $\operatorname{diag}_n(\lambda)$ est permutable avec toute matrice A. La réciproque bien importante pour les applications pratiques est la suivante :

THÉORÈME 3. — *Une matrice de M_n, permutable avec toutes les matrices dans M_n, doit être scalaire.*

DÉMONSTRATION. — Introduisons une matrice E_{ij} dans laquelle l'élément situé à l'intersection de la i-ième ligne et de la j-ième colonne est 1, alors que tous les autres éléments sont nuls. Si $Z = (z_{ij})$ est la matrice dont il s'agit dans le théorème, elle est permutable en particulier avec toutes les E_{ij}:

$$ZE_{ij} = E_{ij}Z, \quad i, j = 1, 2, \dots, n.$$

En multipliant les matrices figurant dans le premier et dans le second membre de cette égalité, on obtient les matrices

$$\underset{\{j\}}{\begin{Vmatrix} 0 & \dots & z_{1i} & \dots & 0 \\ 0 & \dots & z_{2i} & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \dots & z_{ni} & \dots & 0 \end{Vmatrix}} \quad \text{et} \quad \begin{Vmatrix} 0 & 0 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ z_{j1} & z_{j2} & \dots & z_{jn} \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & 0 \end{Vmatrix} \{i\}$$

dont seules j-ième colonne et la i-ième ligne sont respectivement non nulles. Leur comparaison permet de tirer immédiatement les relations $z_{ki} = 0$ pour $k \neq i$, et $z_{ii} = z_{jj}$. Etant donné que i et j varient de 1 à n, on obtient le résultat voulu. ■

On notera encore les relations $\lambda(AB) = (\lambda A) B = A(\lambda B)$ qui découlent immédiatement de la définition du produit d'une matrice par un scalaire ou, si l'on veut, des relations (11) et de l'associativité de la multiplication des matrices.

Pour une matrice donnée $A \in M_n(\mathbb{R})$, on peut essayer de trouver une matrice $B \in M_n(\mathbb{R})$ telle que soit satisfaite la condition

$$AB = E = BA. \tag{12}$$

Si la matrice B existe, cette condition se traduit, en termes de transformations linéaires, par

$$\varphi_A\varphi_B = e = \varphi_B\varphi_A, \tag{12'}$$

qui signifie que $\varphi_B = \varphi_A^{-1}$ est une transformation inverse de φ_A. En vertu du théorème 2 (chap. 1, § 5) φ_A^{-1} existe si, et seulement si, φ_A est une bijection. Ceci étant, φ_A^{-1} est définie de façon unique. Puisque $\varphi_A(0) = 0$, la bijectivité de φ_A signifie en particulier que

$$X \neq 0, \quad X \in \mathbb{R}^n \Rightarrow \varphi_A(X) \neq 0. \qquad (13)$$

Soit maintenant φ_A une transformation linéaire bijective quelconque de $\mathbb{R}^n$. Une transformation inverse φ_A^{-1} existe, mais il n'est pas clair, en général, si elle est linéaire. Pour nous en assurer, introduisons les vecteurs colonnes

$$X = \varphi_A^{-1}(X' + X'') - \varphi_A^{-1}(X') - \varphi_A^{-1}(X''),$$
$$Y = \varphi_A^{-1}(\lambda Y') - \lambda\varphi_A^{-1}(Y')$$

et faisons agir φ_A sur les deux membres de ces égalités. Etant donné que φ_A est linéaire, on obtient

$$\varphi_A(X) = \varphi_A(\varphi_A^{-1}(X' + X'')) - \varphi_A(\varphi_A^{-1}(X')) - \varphi_A(\varphi_A^{-1}(X'')),$$
$$\varphi_A(Y) = \varphi_A(\varphi_A^{-1}(\lambda Y')) - \lambda\varphi_A(\varphi_A^{-1}(Y')).$$

Puisque $\varphi_A\varphi_A^{-1} = e$, on a

$$\varphi_A(X) = e(X' + X'') - e(X') - e(X'') = 0,$$
$$\varphi_A(Y) = e(\lambda Y') - \lambda e(Y') = 0,$$

d'où l'on déduit, conformément à l'implication (13), que X, Y sont des vecteurs nuls. C'est ainsi que sont satisfaites les propriétés (i), (ii) du n° 1 qui définissent les applications linéaires. On a $\varphi_A^{-1} = \varphi_B$, où B est une matrice. Mettons la condition (12') sous la forme $\varphi_{AB} = \varphi_E = \varphi_{BA}$ (voir (8)). En utilisant de nouveau le théorème 1, nous obtenons les égalités (12).

Ainsi, *une matrice inverse de* $A \in M_n(\mathbb{R})$ *existe si, et seulement si, l'application* $\varphi_A: \mathbb{R}^n \to \mathbb{R}^n$ *est bijective. Ceci étant* φ_A^{-1} *est linéaire.* La bijectivité de φ_A est équivalente à la condition que tout vecteur colonne $Y \in \mathbb{R}^n$ s'écrit d'une manière et d'une seule sous la forme (1):

$$Y = \varphi_A(X) = x_1A^{(1)} + x_2A^{(2)} + \ldots + x_nA^{(n)},$$

où $A^{(1)}, A^{(2)}, \ldots, A^{(n)}$ sont les vecteurs colonnes de la matrice A (la surjectivité de φ_A entraîne l'existence de X pour lequel $Y = \varphi_A(X)$, alors que l'injectivité de φ_A donne l'unicité de X: si $Y = \varphi_A(X') = \varphi_A(X'')$, alors $\varphi_A(X' - X'') = \varphi_A(X') - \varphi_A(X'') = 0$, d'où, conformément à (12), $X' - X'' = 0$). Par conséquent, $\mathbb{R}^n$ coïncide avec l'espace des vecteurs colonnes $V_v(A) = \langle A^{(1)}, \ldots, A^{(n)}\rangle$ de la matrice A si bien que rang $A = \dim \mathbb{R}^n = n$.

Si une matrice inverse de A existe, elle est, d'après ce qui précède, unique. On convient de la désigner par le symbole A^{-1}. Dans ce

cas (voir (12′)), on a

$$\varphi_A^{-1} = \varphi_{A^{-1}}. \tag{4}$$

On dit qu'une matrice carrée A est *régulière* (*non dégénérée* ou *non singulière*) si elle admet une matrice inverse A^{-1}. La transformation linéaire correspondante φ_A est appelée, elle aussi, régulière. Dans le cas contraire, la matrice A et la transformation linéaire φ_A sont dites *singulières* (ou *dégénérées*).

Résumons les résultats que nous avons obtenus.

THÉORÈME 4. — *Une matrice carrée A d'ordre n est régulière si, et seulement si, son rang est égal à n. La transformation φ_A^{-1} inverse de φ_A est linéaire et se définit par l'égalité* (14). ■

COROLLAIRE. — *La régularité de φ_A entraîne celle de $\varphi_{A^{-1}}$ et $(A^{-1})^{-1} = A$. Si $A, B, \ldots, C, D$ sont des matrices régulières $n \times n$, le produit $AB \ldots CD$ est aussi régulier et $(AB \ldots CD)^{-1} = D^{-1}C^{-1} \ldots B^{-1}A^{-1}$.*

Pour la démonstration il suffit d'évoquer soit le corollaire du théorème 2 (chap. 1, § 5), soit la symétrie de la condition $AA^{-1} = E = A^{-1}A$. ■

La formule explicite pour A^{-1} sera donnée au chapitre 3. Pour l'instant, nous nous contenterons seulement de remarquer que le calcul de A^{-1} d'une matrice A à éléments numériques, ou le calcul du produit de deux matrices, en appliquant, par exemple, la méthode indiquée à la fin de ce chapitre, exige généralement qu'on exécute un grand nombre d'opérations. En pratique, on a parfois affaire à des matrices d'ordre $n = 100$ et plus. Si A et B sont deux matrices de ce type, il faut, pour calculer $C = AB$, déterminer n^2 éléments c_{ij} d'après la formule (7) (ou (9)), ce qui exige d'effectuer dans chaque cas $2n - 1$ multiplications et additions des nombres. Il faudra exécuter au total $(2n-1)\,n^2$ opérations, c'est-à-dire près de deux millions d'opérations pour $n = 100$. Pour les calculateurs électroniques modernes c'est un problème relativement facile, mais des difficultés réelles surgissent au cas où il s'agit de calculer la *puissance* A^m de la matrice A pour un *exposant* $m \geqslant 1000$. Ici, par définition, $A^m = AA^{m-1}$; la propriété $A^m = A^kA^{m-k}$, $0 \leqslant k \leqslant m$, est une simple conséquence de l'associativité (voir corollaire du théorème 2), comme cela sera montré au chapitre 4, dans un contexte plus général. Pour calculer A^m, on a recours à de diverses méthodes auxiliaires soit basées sur la nature spécifique de la matrice A, soit empruntées à l'algèbre linéaire. Illustrons ce qui vient d'être dit par trois exemples.

EXEMPLE 1. Si

$$A = \operatorname{diag}\{\alpha_1, \ldots, \alpha_n\} = \begin{Vmatrix} \alpha_1 & \ldots & 0 \\ \cdot & \cdot\ \cdot & \cdot \\ 0 & \ldots & \alpha_n \end{Vmatrix},$$

il est évident que

$$A^m = \operatorname{diag}\{\alpha_1^m, \ldots, \alpha_n^m\} = \begin{Vmatrix} \alpha_1^m & \ldots & 0 \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ 0 & \ldots & \alpha_n^m \end{Vmatrix}.$$

EXEMPLE 2. Soit

$$A=\begin{Vmatrix} a & c \\ 0 & b \end{Vmatrix}.$$

Alors, la récurrence sur m montre que

$$A^m=\begin{Vmatrix} a^m & c\,\dfrac{a^m-b^m}{a-b} \\ 0 & b^m \end{Vmatrix},$$

où $\dfrac{a^m-b^m}{a-b}=a^{m-1}+a^{m-2}b+\ldots+ab^{m-2}+b^{m-1}$. En particulier pour $a=b$ on a

$$\begin{Vmatrix} a & c \\ 0 & a \end{Vmatrix}^m=\begin{Vmatrix} a^m & ma^{m-1}c \\ 0 & a^m \end{Vmatrix}.$$

EXEMPLE 3. En raisonnant par récurrence sur m, on s'assure sans peine que la puissance m-ième de la matrice

$$A=\begin{Vmatrix} 0 & 1 \\ 1 & 1 \end{Vmatrix}$$

est de la forme

$$A^m=\begin{Vmatrix} f_{m-1} & f_m \\ f_m & f_{m+1} \end{Vmatrix}, \tag{15}$$

où les entiers $f_0=0$, $f_1=1$, $f_2=1$, $f_3=2$, . . . se déterminent par la relation de récurrence $f_{m+1}=f_m+f_{m-1}$. Ce ne sont rien d'autre que les nombres de Fibonacci (voir exemple 2 en fin du § 3, chap. 1).

Introduisons une matrice

$$B=\begin{Vmatrix} -\dfrac{\lambda_2}{5} & \dfrac{1}{5} \\ -\sqrt{5}\,\lambda_1 & \sqrt{5} \end{Vmatrix}$$

à déterminant 1 (voir chap. 1, § 4), où

$$\lambda_1=\frac{1+\sqrt{5}}{2}, \qquad \lambda_2=\frac{1-\sqrt{5}}{2}.$$

Un simple calcul montre que

$$B^{-1}=\begin{Vmatrix} \sqrt{5} & -\dfrac{1}{5} \\ \sqrt{5}\,\lambda_1 & -\dfrac{\lambda_2}{5} \end{Vmatrix} \quad \text{et} \quad A=B^{-1}\cdot\begin{Vmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{Vmatrix}\cdot B.$$

Or, si A, B, C sont des matrices quelconques de type (n, n), dont B est régulière, qui vérifient la relation $A=B^{-1}CB$, alors on a

$$A^m=B^{-1}CB\cdot B^{-1}CB\cdot B^{-1}CB\ldots B^{-1}CB=B^{-1}C^mB$$

(le produit BB^{-1} est remplacé par E, et l'expression se simplifie). Compte tenu de l'exemple 1 et de la relation (15), on a dans notre cas:

$$\begin{Vmatrix} f_{m-1} & f_m \\ f_m & f_{m+1} \end{Vmatrix} = A^m = B^{-1} \begin{Vmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{Vmatrix}^m B = B^{-1} \begin{Vmatrix} \lambda_1^m & 0 \\ 0 & \lambda_2^m \end{Vmatrix} B =$$

$$= \begin{Vmatrix} \sqrt{5} & -\frac{1}{5} \\ \sqrt{5}\lambda_1 & -\frac{\lambda_2}{5} \end{Vmatrix} \begin{Vmatrix} \lambda_1^m & 0 \\ 0 & \lambda_2^m \end{Vmatrix} B = \begin{Vmatrix} \sqrt{5}\lambda_1^m & -\frac{\lambda_2^m}{5} \\ \sqrt{5}\lambda_1^{m+1} & -\frac{\lambda_2^{m+1}}{5} \end{Vmatrix} \times$$

$$\times \begin{Vmatrix} -\frac{\lambda_2}{5} & \frac{1}{5} \\ -\sqrt{5}\lambda_1 & \sqrt{5} \end{Vmatrix} = \begin{Vmatrix} * & \frac{1}{\sqrt{5}}(\lambda_1^m - \lambda_2^m) \\ * & * \end{Vmatrix}$$

(les astérisques marquent les éléments qui ne nous intéressent pas).

En comparant les éléments des matrices dans le premier et le dernier membre de ces égalités, on obtient pour le nombre de Fibonacci de numéro m la valeur suivante:

$$f_m = \frac{\lambda_1^m - \lambda_2^m}{\sqrt{5}} = \frac{1}{\sqrt{5}} \left\{ \left(\frac{1+\sqrt{5}}{2} \right)^m - \left(\frac{1-\sqrt{5}}{2} \right)^m \right\}.$$

Nous voyons que $f_m \sim \frac{1}{\sqrt{5}} \lambda_1^m$ pour les grandes valeurs de m (progression géométrique), car $\lim_{m\to\infty} \left(\frac{1-\sqrt{5}}{2} \right)^m = 0$.

Nous avons obtenu des règles assez nombreuses qui régissent les opérations sur les matrices carrées d'ordre n. Nous avons en vue les propriétés obtenues à la fin du n° 1, ainsi que l'associativité (corollaire du théorème 2), les relations (10) et le théorème 4. Il importe encore de signaler les *lois* dites *de distributivité*:

$$(A + B)\,C = AC + BC, \quad C\,(A + B) = CA + CB, \qquad (16)$$

où A, B, C sont des matrices arbitraires de M_n ($\mathbb{R}$).

En effet, posant $A = (a_{ij})$, $B = (b_{ij})$, $C = (c_{ij})$, nous obtenons pour tous les $i, j = 1, \ldots, n$ l'égalité (on utilise la distributivité dans $\mathbb{R}$!):

$$\sum_{k=1}^{n} (a_{ik} + b_{ik})\, c_{kj} = \sum_{k=1}^{n} a_{ik}c_{kj} + \sum_{k=1}^{n} b_{ik}c_{kj},$$

dont le premier membre exprime un élément g_{ij} de la matrice $(A + B)\,C$, et le second membre, les éléments respectifs h_{ij} et h'_{ij} des matrices AC et BC. La deuxième loi de distributivité (16) est vérifiée de façon tout à fait analogue. La nécessité de cette loi est due à la non-commutativité de la multiplication dans M_n ($\mathbb{R}$).

Quant aux lois de distributivité

$$(\varphi + \psi)\,\xi = \varphi\xi + \psi\xi, \quad \xi\,(\varphi + \psi) = \xi\varphi + \xi\psi \qquad (16')$$

pour les applications linéaires φ, ψ, ξ de $\mathbb{R}^n$ dans $\mathbb{R}^n$, on peut ne pas les démontrer, vu qu'il existe une bijection entre les applications et les matrices, mais on peut aussi déduire (16) de (16'), étant donné que dans le cas des applications le raisonnement est aussi bien simple

$$((\varphi + \psi)\,\xi)\,(X) = (\varphi + \psi)\,(\xi X) = \varphi\,(\xi X) + \psi\,(\xi X) =$$
$$= (\varphi\xi)\,(X) + (\psi\xi)\,(X) = (\varphi\xi + \psi\xi)\,(X), \quad X \in \mathbb{R}^n.$$

EXERCICES

1. Soient données les applications

a) $[x_1, x_2, \ldots, x_n] \mapsto [x_n, \ldots, x_2, x_1]$;

b) $[x_1, x_2, \ldots, x_n] \mapsto [x_1, x_2^2, \ldots, x_n^n]$;

c) $[x_1, x_2, \ldots, x_n] \mapsto [x_1, x_1 + x_2, \ldots, x_1 + x_2 + \ldots + x_n]$.

Lesquelles d'entre elles sont linéaires?

2. Vérifier que

$$A = \begin{Vmatrix} a & b \\ c & d \end{Vmatrix}, \; ad - bc \neq 0 \Rightarrow A^{-1} = \frac{1}{ad - bc} \begin{Vmatrix} d & -b \\ -c & a \end{Vmatrix}.$$

En particulier, $ad - bc = 1 \Rightarrow A^{-1} = \begin{Vmatrix} d & -b \\ -c & a \end{Vmatrix}$. Est-ce qu'il existe A^{-1} quand $ad - bc = 0$?

3. Démontrer que toute matrice

$$A = \begin{Vmatrix} a & b \\ c & d \end{Vmatrix}$$

vérifie la relation

$$A^2 = (a + d)\,A - (ad - bc)\,E$$

(en d'autres termes, prouver que A est « racine » de l'équation du second degré $x^2 - (a + d)\,x + (ad - bc) = 0$).

4. Etant donné $ad - bc \neq 0$, utiliser la relation de l'exercice 3 pour trouver la matrice inverse A^{-1}.

5. Démontrer que

$$\begin{Vmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{Vmatrix}^m = \begin{Vmatrix} 1 & ma & \frac{m(m-1)}{2}ab + mc \\ 0 & 1 & mb \\ 0 & 0 & 1 \end{Vmatrix}.$$

Déterminer la matrice inverse de la matrice $\begin{Vmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{Vmatrix}$.

6. Vérifier que

$$\begin{Vmatrix} 0 & -1 \\ 1 & -1 \end{Vmatrix}^3 = E.$$

7. Démontrer que, si

$$\begin{Vmatrix} a & b \\ c & d \end{Vmatrix}^m = 0, \quad \text{on a} \quad \begin{Vmatrix} a & b \\ c & d \end{Vmatrix}^2 = 0.$$

8. Dans les applications pratiques, un grand rôle revient à des matrices dites *markoviennes* (ou *stochastiques*) :

$$P = (p_{ij}), \quad p_{ij} \geqslant 0, \quad \sum_{i=1}^{n} p_{ij} = 1, \quad j = 1, 2, \ldots, n.$$

Les applications linéaires φ_P associées à ces matrices sont généralement utilisées pour transformer les vecteurs colonnes spéciaux dits *probabilistes*

$$X = [x_1, \ldots, x_n], \quad x_i \geqslant 0, \quad \sum_{i=1}^{n} x_i = 1.$$

La concordance de ces définitions, qui sont dictées par des problèmes scientifiques, résulte des propositions suivantes qu'il faut démontrer au moins pour $n = 2$:

a) Une matrice $P \in M_n(\mathbb{R})$ est markovienne si, et seulement si, pour tout vecteur probabiliste X le vecteur PX est aussi probabiliste (ici, $PX = \varphi_P(X)$).

b) Si P est une matrice markovienne *positive* ($p_{ij} > 0$, $\forall i, j$), à tout vecteur probabiliste X correspond un vecteur probabiliste *positif* PX (toutes les composantes sont strictement supérieures à zéro).

c) Si P et Q sont des matrices markoviennes, la matrice PQ sera, elle aussi, markovienne. Cela signifie en particulier que toute puissance P^k d'une matrice markovienne est markovienne.

§ 4. Espace des solutions

1. Solutions d'un système linéaire homogène.— Des remarques préliminaires faites au début des §§ 2 et 3 il résulte que le système d'équations linéaires dont la matrice A est de type (m, n) et le vecteur colonne $B \in \mathbb{R}^m$, peut s'écrire tout court sous la forme

$$\varphi_A(X) = B, \tag{1}$$

ou encore

$$AX = B \tag{1'}$$

(le premier membre représente le produit des matrices (m, n) et $(n, 1)$).

Supposons pour un instant que $m = n$ et que la matrice carrée A d'ordre n soit régulière (voir § 3, n° 3). Nous obtenons une solution, d'ailleurs unique, du système (1′) en multipliant à gauche par A^{-1} les deux parties de la relation matricielle : $X = EX = (A^{-1}A)X = A^{-1}(AX) = A^{-1}B$. La matrice A^{-1} n'étant pas donnée d'avance, cette écriture symbolique bien commode des solutions d'un système

carré déterminé ne nous dispense pas des calculs. Néanmoins, l'appareil matriciel développé au § 3, faisons-nous un plaisir de le remarquer, présente au moins un aspect esthétique. Utilisons cet appareil pour la synthèse de toutes les solutions du système (1). A cet effet, considérons d'abord un système homogène associé, lorsque $B = [0, 0, \ldots, 0] = 0$.

Nous appellerons *noyau* d'une application linéaire $\varphi_A : \mathbb{R}^n \to \mathbb{R}^m$ l'ensemble

$$\operatorname{Ker} \varphi_A = \{X \in \mathbb{R}^n \mid \varphi_A (X) = 0\}$$

(la notation Ker provient du mot anglais « *kernel* » traduit en français par « noyau »). Autrement dit, $\operatorname{Ker} \varphi_A$ est l'ensemble des solutions du système homogène de matrice A. En réalité, $\operatorname{Ker} \varphi_A$ est un sous-espace de $\mathbb{R}^n$ (appelé *espace des solutions d'un système linéaire homogène*) ce qu'on a déjà dit au début du § 1 et qu'on déduit sans peine de la propriété de l'application φ_A d'être linéaire :

$$X', X'' \in \operatorname{Ker} \varphi_A \Rightarrow \varphi_A (\alpha X' + \beta X'') = \\ = \alpha\varphi_A (X') + \beta\varphi_A (X'') = 0 \Rightarrow \alpha X' + \beta X'' \in \operatorname{Ker} \varphi_A.$$

A son tour, l'image $\operatorname{Im} \varphi_A$ de l'application φ_A est un sous-espace de $\mathbb{R}^m$: si $B' = \varphi_A (X')$, $B'' = \varphi_A (X'') \in \operatorname{Im} \varphi_A$, on a aussi

$$\alpha B' + \beta B'' = \alpha\varphi_A (X') + \beta\varphi_A (X'') = \varphi_A (\alpha X' + \beta X'') \in \\ \in \operatorname{Im} \varphi_A.$$

La compatibilité du système (1) est équivalente à ce que $B \in \operatorname{Im} \varphi_A$. Soient

$$s = \dim \operatorname{Ker} \varphi_A, \quad r = \dim \operatorname{Im} \varphi_A$$

les dimensions des espaces $\operatorname{Ker} \varphi_A$ et $\operatorname{Im} \varphi_A$. La définition de la dimension, donnée au § 1, entraîne que $s \leqslant n$, $r \leqslant m$. On a aussi $r \leqslant n$, car tout système libre que forment $\varphi_A (X^{(1)}), \ldots, \varphi_A (X^{(k)})$ de $\operatorname{Im} \varphi_A$ ne peut être obtenu qu'à partir du système $X^{(1)}, \ldots, X^{(k)}$ linéairement indépendant dans $\mathbb{R}^n$. Une information plus précise est fournie par le théorème suivant :

THÉORÈME 1. — *On a l'égalité* $r + s = n$. *Le nombre* $r = \dim \operatorname{Im} \varphi_A$ *coïncide avec le rang de la matrice* A (*pour cette raison* r *s'appelle rang de l'application linéaire* φ_A).

DÉMONSTRATION. — Choisissons une base $X^{(1)}, \ldots, X^{(s)}$ du sous-espace $\operatorname{Ker} \varphi_A \subset \mathbb{R}^n$ et complétons-la pour obtenir une base $X^{(1)}, \ldots, X^{(s)}, X^{(s+1)}, \ldots, X^{(n)}$ de l'espace $\mathbb{R}^n$. On peut le faire dans tous les cas comme le montre la démonstration du théorème 2 du § 1 (et l'exercice 5 du § 1). Pour tout vecteur $X = \sum_i \alpha_i X^{(i)} \in \mathbb{R}^n$

on a

$$\varphi_A(X)=\sum_{i=1}^{n}\alpha_i\varphi_A(X^{(i)})=\alpha_{s+1}\varphi_A(X^{(s+1)})+\ldots+\alpha_n\varphi_A(X^{(n)}),$$

de sorte que $\operatorname{Im}\varphi_A=\langle\varphi_A(X^{(s+1)}),\ \ldots,\ \varphi_A(X^{(n)})\rangle$ et $r\leqslant n-s$. Les vecteurs $\varphi_A(X^{(s+1)}),\ \ldots,\ \varphi_A(X^{(n)})$ sont linéairement indépendants vu que $0=\sum\limits_{k\geqslant s+1}\alpha_k\varphi_A(X^{(k)})=\varphi_A\left(\sum\limits_{k\geqslant s+1}\alpha_kX^{(k)}\right)$ implique $\sum\limits_{k\geqslant s+1}\alpha_kX^{(k)}\in\operatorname{Ker}\varphi_A$. Or, en raison du choix de $X^{(s+1)},\ \ldots,\ X^{(n)}$, cela ne peut avoir lieu que pour $\alpha_{s+1}=\ldots=\alpha_n=0$. Cela signifie que $r=n-s$. Par la définition de φ_A, on a pour $X=[x_1,\ \ldots,\ x_n]$: $\varphi_A(X)=x_1A^{(1)}+\ldots+x_nA^{(n)}$, i.e. $\operatorname{Im}\varphi_A=\langle A^{(1)},\ \ldots,\ A^{(n)}\rangle$. Or, la dimension de l'enveloppe linéaire $\langle A^{(1)},\ \ldots,\ A^{(n)}\rangle$ des vecteurs colonnes de la matrice A est justement le rang de A. ■

Nous avons déjà rencontré un cas particulier du théorème 1. Si A est une matrice carrée régulière d'ordre n, le rang maximal possible de A et de φ_A est égal à n.

Pour déterminer la base de l'espace des solutions d'un système linéaire $AX=0$ de rang r, choisissons dans A r vecteurs colonnes de base (le procédé pratique pour ce choix est indiqué au chapitre suivant). Par une permutation des colonnes ou, ce qui revient au même, par un changement de numération des inconnues, on arrive à ce que les vecteurs colonnes de base se représentent par les r premières colonnes $A^{(1)},\ \ldots,\ A^{(r)}$. Tout système de $r+1$ vecteurs colonnes $A^{(1)},\ \ldots,\ A^{(r)},\ A^{(k)},\ k>r$, sera linéairement dépendant et, en vertu du théorème 1 (v) du § 1, on peut écrire un système de relations

$$x_1^{(k)}A^{(1)}+x_2^{(k)}A^{(2)}+\ldots+x_r^{(k)}A^{(r)}+A^{(k)}=0,$$
$$k=r+1,\ r+2,\ \ldots,\ n.$$

Les vecteurs colonnes

$$\begin{aligned}X^{(1)}&=[x_1^{(r+1)},\ x_2^{(r+1)},\ \ldots,\ x_r^{(r+1)},\ 1,\ 0,\ \ldots,\ 0],\\ X^{(2)}&=[x_1^{(r+2)},\ x_2^{(r+2)},\ \ldots,\ x_r^{(r+2)},\ 0,\ 1,\ \ldots,\ 0],\\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ X^{(n-r)}&=[x_1^{(n)},\ x_2^{(n)},\ \ldots,\ x_r^{(n)},\ 0,\ 0,\ \ldots,\ 1]\end{aligned}\qquad(2)$$

dont le nombre est égal à $n-r$, sont évidemment linéairement indépendants (vu la forme spéciale des $n-r$ dernières composantes). Etant solutions du système (1), avec $B=0$, ils constituent, suivant le théorème 1, une base de l'espace $\operatorname{Ker}\varphi_A$ de toutes ses solutions.

Toute base de l'espace des solutions d'un système linéaire homogène $AX=0$ de rang r s'appelle *système fondamental de solutions*. Le système (2) est encore appelé système fondamental *normal*. Suivant le corollaire du théorème 1 du § 2, son rang $s=\dim\operatorname{Im}\varphi_A=$

$= n - r$ est égal au nombre d'inconnues non principales du système linéaire.

Un certain sens « géométrique » des systèmes d'équations linéaires est mis en évidence par l'assertion suivante (dont nous n'aurons pas besoin pour la suite) :

THÉORÈME 2. — *Tout sous-espace $V \subset \mathbb{R}^n$ de dimension s est le sous-espace de solutions d'un système linéaire homogène de rang $r = n - s$.*

DÉMONSTRATION. — Soit $V = \langle A^{(1)}, \ldots, A^{(s)} \rangle$. De même que dans la démonstration du théorème 1, complétons le système libre que forment $A^{(1)}, \ldots, A^{(s)}$ jusqu'à la base $A^{(1)}, \ldots, A^{(s)}, A^{(s+1)}, \ldots, A^{(n)}$ de tout l'espace $\mathbb{R}^n$. Tout vecteur colonne $X = [x_1, \ldots, x_n] \in \mathbb{R}^n$ s'écrit d'une manière et d'une seule sous la forme

$$X = \sum_{j=1}^{n} x'_j A^{(j)} = AX', \tag{3}$$

où A est une matrice carrée d'ordre n formée de vecteurs colonnes $A^{(j)}$ et $X' = [x'_1, x'_2, \ldots, x'_n]$. Puisque les vecteurs colonnes $A^{(1)}, \ldots, A^{(n)}$ sont linéairement indépendants, on a rang $A = n$. En vertu du théorème 4 du § 3, il existe une matrice $A^{-1} = (\bar{a}_{ij})$ inverse de A. On a

$$[x'_1, \ldots, x'_n] = A^{-1}X = [\sum_{j=1}^{n} \bar{a}_{1j}x_j, \ldots, \sum_{j=1}^{n} \bar{a}_{nj}x_j].$$

Le corollaire du théorème 4, § 3, montre que rang $A^{-1} = n$ et, de ce fait, n'importe quelles r lignes de la matrice A^{-1} sont linéairement indépendantes. Par suite, le rang du système homogène

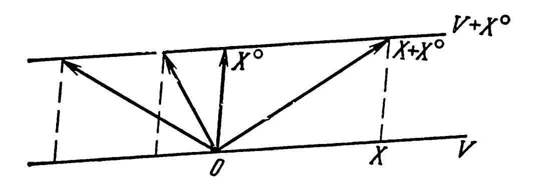
$$\sum_{j=1}^{n} \bar{a}_{kj}x_j = 0, \quad k = s+1, \ldots, n,$$

est $r = n - s$. Or, l'ensemble des solutions de ce système comprend justement ceux des vecteurs colonnes X de la forme (3) dont $x'_{s+1} = 0, \ldots, x'_n = 0$, c'est-à-dire précisément les vecteurs du sous-espace V. ■

2. Variétés linéaires. Solutions d'un système non homogène. — Soient V un sous-espace de $\mathbb{R}^n$ et X° un vecteur fixe de $\mathbb{R}^n$. L'ensemble

$$V + X^\circ = \{X + X^\circ \mid X \in V\} = X^\circ + V$$

est appelé *variété linéaire* de type V et de dimension égale à dim V. La figure géométrique



illustre cette notion d'ailleurs assez claire: $V + X^\circ$ est un espace obtenu de V par une translation de vecteur X°. Le sous-espace V de $\mathbb{R}^n$ est aussi une variété linéaire correspondant à la translation de vecteur $X^\circ = 0$. Deux variétés linéaires de type V coïncident si, et seulement si, elles sont obtenues de V par des translations de vecteurs X' et X'' tels que $X' - X'' \in V$ (nous laissons au lecteur le soin de vérifier cette proposition).

En particulier, si X' est un vecteur arbitraire de la variété linéaire $V + X^\circ$, alors $V + X'$ coïncide avec $V + X^\circ$.

Soient, par exemple, $V = \langle E^{(1)}, E^{(2)}, E^{(3)} \rangle \subset \mathbb{R}^5$, $X^\circ = [0, 0, 1, 1, 0]$, $X' = [0, 0, 0, 1, 0]$. Alors,

$$V + X^\circ = V + X' = \{[x, y, z, 1, 0] \mid x, y, z \in \mathbb{R}\}.$$

Considérons un système d'équations linéaires non homogènes (1). Supposons que le système (1) soit compatible, c'est-à-dire que les rangs des matrices A et $(A \mid B)$ coïncident (théorème 2 du § 2). Soit $X^\circ = [x_1^\circ, \ldots, x_n^\circ]$ une solution fixe quelconque de ce système, telle que $\varphi_A(X^\circ) = B$. Si X' est une autre solution quelconque du système (1), alors $\varphi_A(X' - X^\circ) = \varphi_A(X') - \varphi_A(X^\circ) = B - B = 0$. Cela signifie que la différence $X'' = X' - X^\circ$ de deux solutions du système non homogène (1) est toujours solution du système homogène correspondant et $X' = X'' + X^\circ$. Réciproquement, si $\varphi_A(X) = 0$, on a

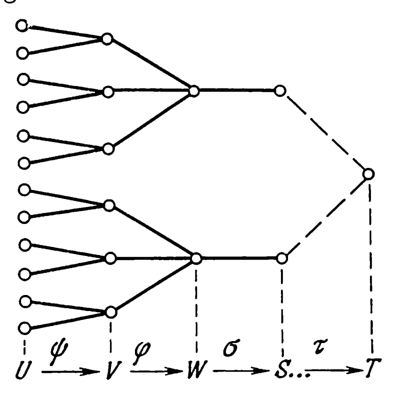
$$\varphi_A(X + X^\circ) = \varphi_A(X) + \varphi_A(X^\circ) = 0 + B = B.$$

Ainsi, nous avons démontré le théorème suivant :

THÉORÈME 3. — *L'ensemble des solutions d'un système linéaire non homogène compatible coïncide avec une variété linéaire de type V, où $V = \operatorname{Ker} \varphi_A$ est le sous-espace vectoriel de solutions du système homogène correspondant.*

3. Rang d'un produit de matrices.— L'action d'un produit $\tau \ldots \sigma\varphi\psi$ d'applications peut être représentée conventionnelle-

ment par le diagramme



qui visualise, dans le cas des applications linéaires des espaces vectoriels, l'implication

$$\varphi\psi\,(U) \subset \varphi\,(V) \Rightarrow \text{rang}\ \varphi\psi \leqslant \text{rang}\ \varphi.$$

Une base de l'espace $\psi\,(U)$ s'applique sur un système de vecteurs contenant la base de l'espace $\varphi\psi\,(U)$, d'où

$$\text{rang}\ \varphi\psi \leqslant \text{rang}\ \psi.$$

Donc,

$$\text{rang}\ \varphi\psi \leqslant \min\,\{\text{rang}\ \varphi,\ \text{rang}\ \psi\}. \qquad (4)$$

Or, rang φ_A = rang A et rang AB = rang φ_{AB} = rang $\varphi_A\varphi_B$, si bien que l'inégalité (4) conduit à l'assertion suivante :

THÉORÈME 4. — *Le rang d'un produit de matrices est au plus égal à celui de chacun des facteurs :*

$$\text{rang}\ AB \leqslant \min\,\{\text{rang}\ A,\ \text{rang}\ B\}. \ \blacksquare \qquad (4')$$

COROLLAIRE 1. — *Si B et C sont deux matrices carrées régulières d'ordres m et n respectivement, et A est une matrice quelconque à m lignes et n colonnes, on a*

$$\text{rang}\ BAC = \text{rang}\ A.$$

DÉMONSTRATION. — D'après le théorème 4 on a

$$\text{rang}\ BAC \leqslant \text{rang}\ BA = \text{rang}\ BA\,(CC^{-1}) =$$
$$= \text{rang}\ (BAC)\,C^{-1} \leqslant \text{rang}\ BAC,$$

d'où rang BAC = rang BA. On établit de même l'égalité rang BA = = rang A. ■

COROLLAIRE 2. — *Si une matrice carrée A d'ordre n possède une inverse unilatère (à droite ou à gauche), elle est régulière.*

DÉMONSTRATION. — Supposons que $AB = E$ pour une matrice B d'ordre n. Puisque rang $E = n$, l'inégalité (4′) se met sous la forme $n \leqslant \min \{\text{rang } A, \text{rang } B\}$, d'où l'on déduit que rang A = rang $B = n$. Or, cette condition équivaut à la régularité de A et B (voir théorème 4 du § 3). On établit de façon analogue la régularité de A dans le cas où il existe une matrice C telle que $CA = E$. ■

En vertu du corollaire 2, l'application linéaire $\varphi_A: \mathbb{R}^n \to \mathbb{R}^n$, inversible à droite ou à gauche, possède une inverse bilatère, ce qui témoigne d'une différence de nature fondamentale qui existe entre les transformations linéaires et les applications en général (voir chap. 1, § 5, exercice 2).

4. Classes de matrices équivalentes. De même qu'au § 3, n° 3, désignons par E_{st} une matrice de type (m, m) dont l'élément situé à l'intersection de la s-ième ligne et de la t-ième colonne est 1, alors que tous les autres éléments sont nuls (de telles matrices sont parfois appelées *unités matricielles*).

Considérons dans $M_m(\mathbb{R})$ les matrices dites *élémentaires*:

(I) $$F_{s,t} = E - E_{ss} - E_{tt} + E_{st} + E_{ts} = \left\| \begin{matrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & 1 & & \\ & & & \ddots & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & 1 & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{matrix} \right\|, \quad s \neq t;$$

(II) $$E_{s,t}(\lambda) = E + \lambda E_{st} = \left\| \begin{matrix} 1 & & \\ & \ddots & \\ \dots & 1 \dots \lambda & \dots \\ & & \ddots \\ & & & 1 \end{matrix} \right\|, \quad s \neq t;$$

(III) $$F_s(\lambda) = E + (\lambda - 1) E_{ss} = \operatorname{diag}\{1, \dots, 1, \lambda, 1, \dots, 1\}, \quad \lambda \neq 0.$$

Soit A une matrice quelconque de type (m, n). On vérifie immédiatement que la matrice $A' = FA$ est obtenue à partir de A en appliquant à ses lignes une transformation élémentaire (t. é.) de type (I) ou (II) suivant qu'on a $F = F_{s,t}$ ou $F = F_{s,t}(\lambda)$. Dans le cas où $F = F_s(\lambda)$ on parlera de la t. é. de type (III) (la multiplication de la s-ième ligne A_s par λ). De façon analogue, la matrice $A'' = AF$ est obtenue en appliquant des transformations élémentaires aux colonnes de la matrice A. Comme nous l'avons vu

au § 2, n° 2 et au § 2, exercice 2, les transformations élémentaires de types (I) et (II), faites sur les lignes et les colonnes, réduisent la matrice A à la forme diagonale. Puisque

$$\begin{Vmatrix} a_1 & & & & & & \\ & a_2 & & & & 0 & \\ & & \ddots & & & & \\ & & & a_r & & & \\ & & & & 0 & & \\ & 0 & & & & \ddots & \\ & & & & & & 0 \end{Vmatrix} = F_1(a_1) F_2(a_2) \ldots F_r(a_r) \begin{Vmatrix} 1 & & & & & & \\ & 1 & & & & 0 & \\ & & \ddots & & & & \\ & & & 1 & & & \\ & & & & 0 & & \\ & 0 & & & & \ddots & \\ & & & & & & 0 \end{Vmatrix},$$

les transformations élémentaires de type (III) permettent d'obtenir à partir de A une matrice de la forme

$$\begin{Vmatrix} E_r & 0 \\ 0 & 0 \end{Vmatrix} \tag{5}$$

(les zéros désignent ici les matrices de types $(r,\ n - r)$, $(m - r,\ r)$ et $(m - r,\ n - r)$). Ainsi

$$P_s P_{s-1} \ldots P_1 A Q_1 Q_2 \ldots Q_t = \begin{Vmatrix} E_r & 0 \\ 0 & 0 \end{Vmatrix}, \tag{6}$$

où P_i (respectivement Q_j) sont des matrices élémentaires d'ordre m (respectivement d'ordre n). Nous avons déjà dit à maintes occasions que les opérations élémentaires sont inversibles, ce qui se concorde avec l'existence de matrices inverses

$$(F_{s,t})^{-1} = F_{s,t}, \quad F_{s,t}(\lambda)^{-1} = \\ = F_{s,t}(-\lambda), \quad F_s(\lambda)^{-1} = F_s(\lambda^{-1}).$$

En vertu du corollaire du théorème 4, § 3, les matrices $P = P_s P_{s-1} \ldots P_1$ et $Q = Q_1 Q_2 \ldots Q_t$ sont aussi inversibles: $P^{-1} = P_1^{-1} \ldots P_{s-1}^{-1} P_s^{-1}$, $Q^{-1} = Q_t^{-1} \ldots Q_2^{-1} Q_1^{-1}$. Remarquons que P_i^{-1}, Q_j^{-1} sont des matrices élémentaires.

On dit que deux matrices A, B de dimensions m et n sont *équivalentes* et on note $A \sim B$ s'il existe des matrices régulières P et Q d'ordre m et n respectivement, telles que $B = PAQ$.

Il est facile de comprendre que $\sim$ est une relation d'équivalence: (i) $A \sim A$ ($P = E_m$, $Q = E_n$); (ii) $A \sim B \Rightarrow B \sim A$, puisque $B = PAQ \Rightarrow A = P^{-1} B Q^{-1}$; (iii) $B = P'AQ'$, $C = P''BQ'' \Rightarrow C = PAQ$, où $P = P''P'$, $Q = Q'Q''$. Suivant les principes généraux (voir chap. 1, § 6), toutes les matrices de dimensions m et n se répartissent par rapport à la relation $\sim$ en classes disjointes de matrices équivalentes. Les rangs des matrices équivalentes étant égaux (voir corollaire 1 du théorème 4), le raisonnement qui nous a conduit à l'égalité (6) montre que les matrices (5) peuvent être

choisies comme représentants des classes d'équivalence. Nous avons ainsi l'assertion suivante :

THÉORÈME 5. — *L'ensemble des matrices de dimensions m et n se répartit en* $P = \min(m, n) + 1$ *classes d'équivalence. Toutes les matrices de rang r appartiennent à une même classe ayant la matrice* (5) *comme représentant.* ■

COROLLAIRE. — *Toute matrice régulière de type* $n \times n$ *s'écrit sous la forme d'un produit de matrices élémentaires.*

En effet, puisque leurs rangs sont égaux à n, toutes les matrices d'ordre n tombent dans une même classe dont le représentant est la matrice unité. La relation (6)

$$P_sP_{s-1} \dots P_1AQ_1Q_2 \dots Q_t = E,$$

mise sous la forme

$$A = P_1^{-1} \dots P_{s-1}^{-1}P_s^{-1}Q_t^{-1} \dots Q_2^{-1}Q_1^{-1}, \tag{7}$$

démontre le théorème. ■

On n'affirme pas que l'écriture de A sous la forme d'un produit de matrices élémentaires est unique, mais le fait lui-même d'existence d'une telle expression est bien utile. En particulier, il peut être mis à profit pour la recherche d'une matrice inverse. En effet, de (7) on tire

$$A^{-1} = Q_1Q_2 \dots Q_tP_sP_{s-1} \dots P_1 = QP.$$

Puisqu'à chacune des matrices P_i et Q_j dans la formule (6) correspond une transformation élémentaire, la suite de transformations

$$E\wr (A) \to P_1\wr (P_1A) \to \dots \to P_s \dots P_1 \wr (P_s \dots P_1A) \to$$

$$\to P_s \dots P_1\wr (P_s \dots P_1AQ_1) \wr Q_1 \to \dots$$

$$\dots \to P_s \dots P_1\wr (P_s \dots P_1AQ_1 \dots Q_t) \wr Q_1 \dots Q_t$$

est pratiquement réalisable bien que le nombre $s + t$ de toutes les transformations puisse être très grand. Nous avons encadré, pour mémoire, par des traits verticaux ondulés $\wr$ les produits des matrices qui nous intéressent, c'est-à-dire les résultats des transformations élémentaires faites sur les lignes (à gauche) et sur les colonnes (à droite) de la matrice unité. Pour $r < n$, nous concluons que A est une matrice singulière et ne possède pas de matrice inverse. Pour $r = n$, il ne nous reste qu'à multiplier Q et P pour obtenir A^{-1}. Remarquons que l'ordre des transformations faites sur les lignes et sur les colonnes peut être changé.

Considérons deux exemples.
Pour la matrice

$$A=\begin{Vmatrix}1&2&3\\4&5&6\\7&8&9\end{Vmatrix}$$

on a

$$E\begin{Vmatrix}1&2&3\\4&5&6\\7&8&9\end{Vmatrix}\to F_{2,1}(-4)\begin{Vmatrix}1&2&3\\0&-3&-6\\7&8&9\end{Vmatrix}\to$$

$$\to F_{3,1}(-7)\,F_{2,1}(-4)\begin{Vmatrix}1&2&3\\0&-3&-6\\0&-6&-12\end{Vmatrix}\to$$

$$\to F_{3,2}(-2)\cdot F_{3,1}(-7)\cdot F_{2,1}(-4)\begin{Vmatrix}1&2&3\\0&-3&-6\\0&0&0\end{Vmatrix}.$$

Vu que la matrice obtenue est de rang 2, on a rang $A = 2$. Par suite, A est une matrice singulière.

La suite

$$E\begin{Vmatrix}0&0&0&1\\1&0&0&0\\0&1&0&0\\0&0&1&0\end{Vmatrix}\to F_{1,2}\begin{Vmatrix}1&0&0&0\\0&0&0&1\\0&1&0&0\\0&0&1&0\end{Vmatrix}\to F_{2,3}F_{1,2}\begin{Vmatrix}1&0&0&0\\0&1&0&0\\0&0&0&1\\0&0&1&0\end{Vmatrix}\to$$

$$\to F_{3,4}F_{2,3}F_{1,2}\begin{Vmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{Vmatrix}$$

entraîne

$$\begin{Vmatrix}0&0&0&1\\1&0&0&0\\0&1&0&0\\0&0&1&0\end{Vmatrix}^{-1}=F_{3,4}F_{2,3}F_{1,2}=\begin{Vmatrix}0&1&0&0\\0&0&1&0\\0&0&0&1\\1&0&0&0\end{Vmatrix}$$

(en pratique, les produits $F_{2,3}F_{1,2}$, $F_{3,4}\,F_{1,2}$ seraient écrits de suite sous forme explicite).

Le calcul de la matrice inverse par cette nouvelle méthode que l'on appelle parfois *(P, Q)-réduction de la matrice à la forme canonique* (5), est assez commode, mais il est encore prématuré de parler de ses avantages et inconvénients en se basant sur les exemples simples que nous venons de considérer : nous n'avons même pas utilisé toutes les matrices $F_{s,t}$, $F_{s,t}(\lambda)$, $F_s(\lambda)$.

EXERCICES

1. Donner les règles opératorielles sur les matrices transposées (voir § 2, exercice 1):

$$^t(A+B) = {}^tA + {}^tB;$$
$$^t(AB) = {}^tB \cdot {}^tA.$$

2. Démontrer par un raisonnement direct sur les matrices que rang $AB \leqslant \min\{\text{rang } A, \text{ rang } B\}$. (Indication. On remarquera que si les vecteurs colonnes de base de la matrice B sont ceux d'indices $j_1, \ldots, j_r$, tous les vecteurs colonnes de la matrice AB s'expriment linéairement en fonction des vecteurs colonnes $(AB)^{(k)}$, $k = j_1, \ldots, j_r$. La même remarque est valable pour la matrice transposée $^t(AB) = {}^tB{}^tA$.)

3. Démontrer l'*inégalité de Sylvester*

$$\dim \operatorname{Ker} \varphi\psi \leqslant \dim \operatorname{Ker} \varphi + \dim \operatorname{Ker} \psi$$

pour deux applications linéaires arbitraires $\mathbb{R}^n \xrightarrow{\psi} \mathbb{R}^m \xrightarrow{\varphi} \mathbb{R}^l$. (Indication. Considérer la restriction $\bar{\varphi} = \varphi \mid_V$ de l'application φ à un sous-espace $V \subset \mathbb{R}^m$. Il est évident que $\operatorname{Ker} \bar{\varphi} \subset \operatorname{Ker} \varphi$. Par conséquent, d'après le théorème 1 (nous savons déjà que V peut être interprété comme $\mathbb{R}^k$, $k \leqslant m$, et de ce fait, le théorème 1 est applicable), $\dim V - \text{rang } \bar{\varphi} = \dim \operatorname{Ker} \bar{\varphi} \leqslant \dim \operatorname{Ker} \varphi$, d'où $\dim V - \dim \varphi(V) \leqslant \dim \operatorname{Ker} \varphi$. En posant $V = \psi(\mathbb{R}^n) = \operatorname{Im} \psi$, on obtient finalement $\dim \operatorname{Ker} \varphi\psi = n - \text{rang } \varphi\psi = (n - \text{rang } \psi) + (\dim V - \text{rang } \bar{\varphi}) \leqslant \leqslant \dim \operatorname{Ker} \psi + \dim \operatorname{Ker} \varphi$.)

4. Démontrer que toute application linéaire $\varphi: \mathbb{R}^n \to \mathbb{R}^m$ de rang r s'écrit sous forme de la somme $\varphi = \varphi_1 + \ldots + \varphi_r$ des applications φ_i de rang 1.

5. Calculer le rang de la matrice

$$A = \begin{Vmatrix} x_1y_1 & x_1y_2 & \ldots & x_1y_n \\ x_2y_1 & x_2y_2 & \ldots & x_2y_n \\ \cdot & \cdot & \cdot & \cdot \\ x_ny_1 & x_ny_2 & \ldots & x_ny_n \end{Vmatrix}.$$

Indication. Montrer que $A = [x_1, \ldots, x_n] (y_1, \ldots, y_n)$.)

CHAPITRE 3

DÉTERMINANTS

Les formules (3) et (9) obtenues au § 4 du chapitre 1 pour les solutions des systèmes linéaires carrés d'ordres $n = 2, 3$ incitent à penser que des formules analogues existent aussi pour n quelconque. Tout compte fait, il s'agit dans chacune des formules mentionnées de donner une interprétation correcte au numérateur et au dénominateur. Nous les considérerons comme valeurs d'une certaine fonction « universelle » $\det : M_n(\mathbb{R}) \to \mathbb{R}$ de l'ensemble des matrices carrées d'ordre n dans $\mathbb{R}$. La construction effective de la fonction det (du déterminant) permettra aussi de donner la réponse à plusieurs autres questions relatives aux matrices, qui ont été soulevées au chapitre 2. En réalité, le rôle que la théorie des déterminants joue en mathématiques, dépasse largement le sujet abordé, et chacune des applications de cette théorie suggère sa propre voie à suivre pour la construire. L'une des approches les plus naturelles du problème des déterminants est l'approche géométrique basée sur l'analogie « déterminants des matrices — volumes des figures à plusieurs dimensions » (voir chap. 1, § 4, exercice 3) et sur les n-formes extérieures. Une telle manière de procéder exigeant un peu plus de géométrie, nous arrêterons notre choix sur la voie « analytique » *).

§ 1. Déterminants: construction et propriétés essentielles

1. Construction par récurrence.— Convenons de considérer que le déterminant de la matrice (a_{11}) de type 1×1 est égal au nombre a_{11}. Les déterminants des matrices 2×2 et 3×3 sont définis res-

*) Il existe plusieurs méthodes analytiques permettant d'exposer la théorie des déterminants. Dans ce chapitre, de même qu'au chapitre 1, § 4, nous nous laissons guider par les conférences de I. R. Chafarévitch (Editions de l'Université de Moscou, 1971), en estimant qu'un exercice de plus sur la méthode de récurrence est utile en soi-même. En tout cas, les procédés pratiquement importants de calcul des déterminants de matrices s'obtiennent assez facilement, bien que l'exposé basé sur la formule du « développement complet du déterminant » (voir chap. 4, § 3) soit, peut-être, un peu plus simple (A. G. Kurosh, « Cours d'algèbre supérieure »).

pectivement par les formules (2) et (8) du chapitre 1, § 4. Dans le dernier cas, les déterminants des matrices de type 2×2 sont laissés exprès sous une forme « implicite ». Nous soulignons par là la base de récurrence que nous nous proposons d'utiliser pour construire les déterminants des matrices de type $n \times n$.

Supposons déjà introduits les déterminants des matrices d'ordres 1, 2, . . ., $n - 1$. Appelons *déterminant de la matrice* $A = (a_{ij})_1^n$ la quantité

$$D = a_{11}D_1 - a_{21}D_2 + \ldots + (-1)^{n-1}a_{n1}D_n, \qquad (1')$$

où D_k est le déterminant de la matrice d'ordre $n - 1$:

$$\left\| \begin{matrix} a_{12} & \ldots & a_{1n} \\ \cdot & \cdot & \cdot \\ a_{k-1,2} & \ldots & a_{k-1,n} \\ a_{k+1,2} & \ldots & a_{k+1,n} \\ \cdot & \cdot & \cdot \\ a_{n2} & \ldots & a_{nn} \end{matrix} \right\|$$

obtenue à partir de A en supprimant la première colonne et la k-ième ligne.

Il est facile de s'assurer que pour $n = 2, 3$ l'expression (1′) est en accord avec les expressions (2), (8) du chapitre 1, § 4. Le déterminant de la matrice $A = (a_{ij})$ est désigné par les symboles $|A|$, $|a_{ij}|_1^n$ ou bien, le plus souvent, par det A. Les barres verticales sont surtout utilisées dans le cas où la matrice A s'écrit sous sa forme explicite.

Si, dans la matrice A, on supprime la i-ième ligne et la j-ième colonne, en laissant inchangée la disposition des autres éléments, on obtient une matrice carrée d'ordre $(n - 1)$. Le déterminant de cette dernière se note M_{ij} et s'appelle le *mineur* de la matrice A correspondant à l'élément a_{ij}.

Avec les nouvelles notations, la formule (1′) prend la forme

$$\det A = a_{11}M_{11} - a_{21}M_{21} + \ldots + (-1)^{n-1}a_{n1}M_{n1}, \qquad (1)$$

qui se traduit de la manière suivante : *le déterminant d'une matrice carrée d'ordre n est égal à la somme algébrique des produits des éléments de la première colonne par les mineurs correspondants, ces produits étant affectés de signes alternés.*

Si, au lieu de la première colonne, on prend la k-ième et l'on remplace les mineurs M_{i1} par les mineurs M_{ik}, on obtient, comme nous le verrons plus loin, une expression qui ne diffère de det A que par le signe.

De même qu'au chapitre 2, nous désignerons dans ce qui suit par les symboles

$$A_i = (a_{i1}, a_{i2}, \ldots, a_{in}), \quad i = 1, 2, \ldots, n,$$
$$A^{(j)} = [a_{1j}, a_{2j}, \ldots, a_{nj}], \quad j = 1, 2, \ldots, n,$$

respectivement la i-ième ligne et la j-ième colonne de la matrice $A = (a_{ij})$. Quant à la matrice A, elle sera représentée soit par la réunion de ses lignes

$$A = [A_1, A_2, \ldots, A_n]$$

(colonne de lignes), soit par la réunion de ses colonnes:

$$A = (A^{(1)}, A^{(2)}, \ldots, A^{(n)})$$

(ligne de colonnes). Convenons d'appeler par la suite *lignes* et *colonnes du déterminant* $| a_{ij} |$ *d'ordre* n les lignes et les colonnes de la matrice A de type (n, n).

Par définition, $| \ \ | = \det$ est une fonction qui associe à une matrice carrée A un nombre $| A | = \det A$. Proposons-nous d'étudier le comportement de cette fonction en cas de changement des lignes ou des colonnes de la matrice A, considérés en tant qu'éléments (vecteurs) de l'espace vectoriel $\mathbb{R}^n$. Si l'on préfère, $\det A$ signifie pour nous une désignation abrégée (dans l'esprit du chap. 1, § 5, nº 2) de la fonction

$$\det [A_1, \ldots, A_n] \text{ ou } \det (A^{(1)}, \ldots, A^{(n)})$$

de n variables qui sont les vecteurs de $\mathbb{R}^n$.

Nous dirons qu'une fonction arbitraire $\mathscr{D}: [A_1, \ldots, A_n] \mapsto \mapsto \mathscr{D}(A_1, \ldots, A_n)$ est *multilinéaire* si elle est linéaire par rapport à chaque argument A_i, c'est-à-dire, si

$$\mathscr{D}(A_1, \ldots, \alpha A'_i + \beta A''_i, \ldots, A_n) =$$
$$= \alpha \mathscr{D}(A_1, \ldots, A'_i, \ldots, A_n) + \beta \mathscr{D}(A_1, \ldots, A''_i, \ldots, A_n)$$

(comparer avec chap. 2, § 3, nº 1). La même fonction $\mathscr{D}$ est dite *symétrique gauche* si

$$\mathscr{D}(A_1, \ldots, A_i, A_{i+1}, \ldots, A_n) =$$
$$= - \mathscr{D}(A_1, \ldots, A_{i+1}, A_i, \ldots, A_n), \ 1 \leqslant i \leqslant n-1. \quad (2)$$

REMARQUE 1. — On déduit de la définition des fonctions linéaires (voir chap. 2, § 3, (4)) que la fonction $\mathscr{D}$ est multilinéaire si, et seulement si, pour $A_1, \ldots, A_{i-1}, A_{i+1}, \ldots, A_n$ fixes et pour $A_i = X = (x_1, \ldots, x_n)$ on a

$$\mathscr{D}(A_1, \ldots, A_n) = \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n,$$

où $\alpha_1, \ldots, \alpha_n$ sont des scalaires indépendants de $x_1, \ldots, x_n$.

REMARQUE 2. — Pour qu'une multilinéaire $\mathscr{D}$ soit symétrique gauche, il faut et il suffit que

$$\mathscr{D}(A_1, \ldots, A_{i-1}, X, X, A_{i+2}, \ldots, A_n) = 0,$$
$$1 \leqslant i \leqslant n-1. \tag{2'}$$

En effet, posant $A_i = A_{i+1} = X$ dans (2), nous retrouvons (2'). Réciproquement, pour $X = A_i + A_{i+1}$ on déduit de (2'), vu que $\mathscr{D}$ est multilinéaire, la relation

$$\mathscr{D}(\ldots, A_i, A_i, \ldots) + \mathscr{D}(\ldots, A_{i+1}, A_{i+1}, \ldots) +$$
$$+ \mathscr{D}(\ldots, A_i, A_{i+1}, \ldots) + \mathscr{D}(\ldots, A_{i+1}, A_i, \ldots) =$$
$$= \mathscr{D}(\ldots, A_i + A_{i+1}, A_i + A_{i+1}, \ldots) = 0.$$

Les deux premiers termes étant nuls (poser dans (2') respectivement $X = A_i$ et $X = A_{i+1}$), la somme de deux derniers termes est égale à zéro, ce qui n'est qu'une autre écriture de la relation (2).

Les mêmes définitions et les mêmes remarques sont valables pour la fonction $\mathscr{D}(A^{(1)}, \ldots, A^{(n)})$ des vecteurs colonnes. De plus, la condition (2), présente dans la définition de la fonction symétrique gauche $\mathscr{D}$, est applicable à toute fonction $\mathscr{D}: M^n \to \mathbb{R}$, où M^n est la puissance cartésienne d'un certain ensemble M. Par la suite, nous aurons besoin de l'assertion suivante:

LEMME 1. — *Si l'on permute deux arguments quelconques, la fonction symétrique gauche change de signe.*

DÉMONSTRATION. — Soient permutés les i-ième et j-ième arguments, $i < j$. Raisonnons par récurrence sur le nombre $k = j - i - 1$ d'arguments situés entre les deux arguments à permuter. Pour $k = 0$ l'assertion du lemme coïncide avec la définition de la fonction symétrique gauche. Supposons le lemme vrai pour tous les $j - i - 1 < k$. Alors

$$\mathscr{D}(\ldots, X_i, X_{i+1}, \ldots, X_{j-1}, X_j, \ldots) =$$
$$= -\mathscr{D}(\ldots, X_{i+1}, X_i, \ldots, X_{j-1}, X_j, \ldots) =$$
$$= \mathscr{D}(\ldots, X_{i+1}, X_j, \ldots, X_{j-1}, X_i, \ldots) =$$
$$= -\mathscr{D}(\ldots, X_j, X_{i+1}, \ldots, X_{j-1}, X_i, \ldots). \blacksquare$$

2. Propriétés fondamentales des déterminants.— La définition du déterminant que nous avons introduite, est encore peu efficace. Il nous faut établir toute une série de propriétés des déterminants (ou plus exactement, de la fonction det) qui soient commodes tant au point de vue de la théorie qu'à celui du calcul.

La relation triviale $\det(a + b) = \det a + \det b$, valable pour les déterminants du premier ordre, peut conduire à une conclusion qu'elle est aussi vraie pour les déterminants d'ordre n (indiquer un

exemple pour $n = 2$). Le cas de $n = 2$ suggère une interprétation plus exacte de la relation considérée :

$$\begin{vmatrix} \alpha x_1' + \beta x_1'' & \alpha x_2' + \beta x_2'' \\ a_{21} & a_{22} \end{vmatrix} = (\alpha x_1' + \beta x_1'')\, a_{22} - (\alpha x_2' + \beta x_2'')\, a_{21} =$$
$$= \alpha\,(x_1' a_{22} - x_2' a_{21}) + \beta\,(x_1'' a_{22} - x_2'' a_{21}) =$$
$$= \alpha \begin{vmatrix} x_1' & x_2' \\ a_{21} & a_{22} \end{vmatrix} + \beta \begin{vmatrix} x_1'' & x_2'' \\ a_{21} & a_{22} \end{vmatrix}.$$

Nous remarquons encore que

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = - \begin{vmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{vmatrix}, \quad \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1.$$

Ainsi, il y a tout lieu d'énoncer le théorème suivant :

THÉORÈME 1. — *La fonction* $A \mapsto \det A$ *sur l'ensemble* $M_n(\mathbb{R})$ *possède les propriétés suivantes :*

D1. det A *est une fonction multilinéaire des lignes de la matrice* A, *c'est-à-dire le déterminant de la matrice est une fonction linéaire des éléments de toute ligne* A_i.

D2. det A *est une fonction symétrique gauche des lignes de la matrice* A (*en d'autres termes, le déterminant est nul si ses deux lignes voisines quelconques coïncident*).

D3. det $E = 1$.

DÉMONSTRATION. — Raisonnons par récurrence sur n. Pour $n = 1, 2$, les propriétés D1 à D3 sont vérifiées. Supposons que tous les déterminants d'ordre $< n$ jouissent de ces propriétés. Démontrons D1 à D3 pour les déterminants d'ordre n, en partant de la formule (1). Commençons par la propriété D3.

D3. Si

$$A = E = \begin{Vmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & 1 \end{Vmatrix},$$

on aura dans la formule (1) $a_{i1} = 0$ pour $i \neq 1$ et $a_{11} = 1$, par suite, det $E = M_{11}$. Le déterminant M_{11} a la même structure que celle de det E, mais son ordre est égal à $n - 1$. Par hypothèse de récurrence nous pouvons poser $M_{11} = 1$ et donc det $E = 1$.

Quant aux propriétés D1, D2, nous allons les démontrer dans une situation un peu générale, décrite par le lemme suivant :

LEMME 2. — *Soit* $\mathscr{D}_j \colon M_n(\mathbb{R}) \to \mathbb{R}$ *une fonction définie par la formule*

$$\mathscr{D}_j(A) = a_{1j} M_{1j} - a_{2j} M_{2j} + \ldots + (-1)^{n-1} a_{nj} M_{nj} \qquad (3)$$

(par hypothèse de récurrence nous connaissons tous les déterminants M_{kj} d'ordre $n-1$, si bien que la fonction $\mathscr{D}_j$ est définie de façon correcte).

Alors, les assertions suivantes sont vraies :

D_j1. *$\mathscr{D}_j$ est une fonction multilinéaire des lignes de la matrice A ;*

D_j2. *$\mathscr{D}_j$ est une fonction symétrique gauche des lignes de la matrice A*

DÉMONSTRATION. — D_j1. Pour souligner la nature variable des éléments de la i-ième ligne, posons $x_s = a_{is}$, $s = 1, \ldots, n$:

$$A = \begin{Vmatrix} a_{11} & \ldots & a_{1j} & \ldots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{i-1,1} & \ldots & a_{i-1,j} & \ldots & a_{i-1,n} \\ x_1 & \ldots & x_j & \ldots & x_n \\ a_{i+1,1} & \ldots & a_{i+1,j} & \ldots & a_{i+1,n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \ldots & a_{nj} & \ldots & a_{nn} \end{Vmatrix}.$$

Le mineur M_{ij} ne dépend pas de $x_1, \ldots, x_n$, si bien que $\alpha_j = (-1)^{i-1}M_{ij}$ est une constante. Tout autre mineur M_{kj}, $k \neq i$, contient parmi ses lignes la ligne $(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)$, toutes les autres étant constantes. Par hypothèse de récurrence, M_{kj} est une fonction linéaire des variables $x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n$, c'est-à-dire on a, en vertu de la remarque 1,

$$M_{kj} = \sum_{s \neq j} \alpha_{ks} x_s, \qquad k \neq i.$$

En posant maintenant $\alpha_s = \sum_{k \neq i} (-1)^{k-1} \alpha_{ks} a_{kj}$, $s \neq j$, nous obtenons l'expression

$$\mathscr{D}_j(A) = \sum_{k=1}^{n} (-1)^{k-1} a_{kj} M_{kj} =$$
$$= \alpha_j x_j + \sum_{k \neq i} (-1)^{k-1} a_{kj} \sum_{s \neq j} \alpha_{ks} x_s =$$
$$= \alpha_j x_j + \sum_{s \neq j} \Big(\sum_{k \neq i} (-1)^{k-1} \alpha_{ks} a_{kj} \Big) x_s = \sum_{s=1}^{n} \alpha_s x_s,$$

qui signifie que $\mathscr{D}_j(A)$ est une fonction linéaire des éléments $x_1, \ldots, x_n$ de la i-ième ligne de la matrice A.

D_j2. Conformément à la remarque 2 du nº 1 il serait plus commode de démontrer l'égalité $\mathcal{D}_j(A) = 0$ pour la matrice

$$A = \begin{Vmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ x_1 & \dots & x_j & \dots & x_n \\ x_1 & \dots & x_j & \dots & x_n \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \dots & a_{nj} & \dots & a_{nn} \end{Vmatrix}$$

contenant deux lignes identiques $A_i = A_{i+1} = (x_1, \dots, x_j, \dots \dots, x_n)$. Le mineur M_{kj}, $k \neq i, i+1$, contient, lui aussi, deux lignes consécutives identiques $(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)$ de longueur $n-1$. C'est pourquoi, par hypothèse de récurrence, on a $M_{kj} = 0$, $k \neq i$, $i+1$. La formule (3) se met sous la forme

$$\mathcal{D}_j(A) = (-1)^{i-1} x_j M_{ij} + (-1)^i x_j M_{i+1,\,j}.$$

Or, il est évident que $M_{ij} = M_{i+1,\,j}$. Par conséquent

$$\mathcal{D}_j(A) = (-1)^{i-1} x_j (M_{ij} - M_{i+1,\,j}) = 0. \ ■$$

En posant $j = 1$ dans la formule (3) et en comparant l'expression obtenue à la formule (1), nous obtenons l'égalité

$$\mathcal{D}_1(A) = \det A. \qquad (4)$$

Par conséquent, les propriétés D1, D2 des déterminants sont contenues dans l'assertion du lemme. Le théorème 1 est démontré. ■

Mettons la propriété D1 sous une forme plus détaillée:

D1′. $\det[A_1, \dots, \lambda A_i, \dots, A_n] = \lambda \det[A_1, \dots, A_i, \dots, A_n]$,

cela veut dire que, *lorsqu'une ligne A_i du déterminant est multipliée par λ, le déterminant lui-même se trouve multiplié par λ. En particulier, si l'on multiplie toutes les lignes par λ, on obtient*

$$\det \lambda A = \lambda^n \det A.$$

D1″. *Si, pour un certain i, tous les éléments de A_i sont de la forme $a_{ij} = a'_j + a''_j$, alors* $\det A = \det A' + \det A''$, *où $A'_j = A''_j = A_j$ pour $j \neq i$, et $A'_i = (a'_1, \dots, a'_n)$, $A''_i = (a''_1, \dots, a''_n)$.*

Du théorème 1 il résulte encore quelques assertions simples que nous énoncerons comme propriétés des déterminants, mais démontrerons pour toute fonction $\mathcal{D}_j$ définie par la formule (3). Le passage aux déterminants est dans ce cas assuré par l'égalité (4).

D4. *Un déterminant contenant une ligne nulle est égal à zéro.*

Soit, par exemple, $A_i = (0, 0, \ldots, 0)$. Alors, on a aussi $2A_i = (0, \ldots, 0)$. Donc, d'après D_j1 :

$$\mathcal{D}_j(A) = \mathcal{D}_j(A_1, \ldots, A_i, \ldots, A_n) = \mathcal{D}_j(A_1, \ldots, 2A_i, \ldots, A_n) = 2\mathcal{D}_j(A_1, \ldots, A_i, \ldots, A_n) = 2\mathcal{D}_j(A),$$

d'où $\mathcal{D}_j(A) = 0$. ■

D5. *L'échange de deux lignes quelconques (et non seulement des lignes voisines) inverse le signe du déterminant.*

Pour toute fonction $\mathcal{D}_j(A)$ cette propriété s'ensuit de D_j2 et du lemme 1. ■

D6. *Si dans une matrice carrée A deux lignes sont identiques, son déterminant est nul.*

Prenons de nouveau une fonction arbitraire $\mathcal{D}_j(A)$. En permutant deux lignes identiques A_s et A_t dans A, nous obtenons la même matrice A. D'autre part, d'après la propriété D5 (plus exactement, d'après la propriété D_j5 pour $\mathcal{D}_j$), $\mathcal{D}_j(A)$ changera de signe. Ainsi, $\mathcal{D}_j(A) = -\mathcal{D}_j(A)$, d'où $2\mathcal{D}_j(A) = 0$ et $\mathcal{D}_j(A) = 0$. ■

D7. *Un déterminant ne change pas si l'on applique à ses lignes les transformations élémentaires de type* (II).

Il suffit de considérer le cas de l'application d'une seule transformation élémentaire. Soit A' une matrice obtenue de la matrice A en ajoutant à la s-ième ligne de A sa t-ième ligne multipliée par λ. En vertu des propriétés D1 et D6 (plus exactement, D_j1 et D_j6 pour $\mathcal{D}_j$), on a alors

$$\mathcal{D}_j(A') = \mathcal{D}_j(A_1, \ldots, A_s + \lambda A_t, \ldots, A_n) = \mathcal{D}_j(\ldots, A_s, \ldots) + \lambda\mathcal{D}_j(\ldots, A_t, \ldots, A_t, \ldots) = \mathcal{D}_j(A_1, \ldots, A_n) = \mathcal{D}_j(A).$$ ■

Les propriétés que nous venons de démontrer permettent de calculer, de façon relativement simple, le déterminant d'ordre n. L'une des méthodes permettant ce calcul est la suivante. En appliquant des transformations élémentaires, il convient de réduire la matrice $A = (a_{ij})$ à sa forme triangulaire (voir chap. 1, § 3). Soit

$$\bar{A} = \left\| \begin{matrix} \bar{a}_{11} & \bar{a}_{12} & \ldots & \bar{a}_{1n} \\ 0 & \bar{a}_{22} & \ldots & \bar{a}_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \ldots & \bar{a}_{nn} \end{matrix} \right\| \qquad (5)$$

la matrice obtenue. Supposons que pour cette réduction on ait fait q transformations élémentaires de type (I) et un certain nombre de transformations de type (II). Puisque ces dernières ne changent pas le déterminant (la propriété D7) et chaque transformation de type

(I) le multiplie par (—1), on a $\det \bar{A} = (-1)^q \det A$ *). Nous allons démontrer que

$$\det \bar{A} = \bar{a}_{11}\bar{a}_{22} \ldots \bar{a}_{nn}.$$

Dans ce cas on aura

$$\det A = (-1)^q \bar{a}_{11}\bar{a}_{22} \ldots \bar{a}_{nn}. \quad (6)$$

Ce sera justement l'une des formules pour le calcul de det A.

Démontrons cette formule pour $\det \bar{A}$, en raisonnant par récurrence sur n. Puisque $\bar{a}_{21} = \ldots = \bar{a}_{n1} = 0$, on a, d'après (1), $\det \bar{A} = \bar{a}_{11} \bar{M}_{11}$, où

$$\bar{M}_{11} = \begin{Vmatrix} \bar{a}_{22} & \bar{a}_{23} & \ldots & \bar{a}_{2n} \\ 0 & \bar{a}_{33} & \ldots & \bar{a}_{3n} \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \bar{a}_{nn} \end{Vmatrix}$$

est un déterminant d'ordre $n - 1$. Par hypothèse de récurrence, $\bar{M}_{11} = \bar{a}_{22}\bar{a}_{33} \ldots \bar{a}_{nn}$. Par suite, $\det \bar{A} = \bar{a}_{11}\bar{M}_{11} = \bar{a}_{11} \bar{a}_{22} \ldots \bar{a}_{nn}$.

Maintenant, en nous appuyant sur la formule (6), nous allons établir un fait important relatif au rôle que jouent les propriétés D1 à D3 des déterminants. A savoir, on peut énoncer le théorème suivant:

Théorème 2. — *Soit $\mathcal{D}: M_n(\mathbb{R}) \to \mathbb{R}$ une fonction ayant les propriétés suivantes:*

(i) *$\mathcal{D}(A)$ est une fonction linéaire des éléments de chaque ligne de la matrice $A \in M_n(\mathbb{R})$;*

(ii) *l'échange de deux lignes voisines remplace $\mathcal{D}(A)$ par son opposé (autrement dit, $\mathcal{D}(A)$ est une fonction symétrique gauche multilinéaire des lignes de la matrice).*

Alors, il existe une constante ρ indépendante de A et telle que l'on a

$$\mathcal{D}(A) = \rho \cdot \det A.$$

Le nombre ρ est déterminé par la relation $\rho = \mathcal{D}(E)$, où E est la matrice unité.

Démonstration. — Suivant le lemme 1, la fonction $\mathcal{D}(A)$ change de signe lors de la permutation de n'importe quelles deux lignes, c'est-à-dire lors de toute transformation élémentaire de type (I). Un raisonnement tout à fait analogue à celui que nous avons développé pour démontrer la propriété D7 montre que $\mathcal{D}(A)$ ne change

---

*) Il convient de signaler que nous aurions pu réduire la matrice A à la forme triangulaire à l'aide des seules transformations élémentaires (des lignes) de type (II) qui ne changent pas le signe du déterminant. Dans ce cas, nous n'aurions pas besoin d'utiliser pour la démonstration le facteur $(-1)^q$.

pas si les lignes de la matrice A sont soumises à une transformation élémentaire de type (II).

Réduisons la matrice A, au moyen des transformations élémentaires, à la forme triangulaire (5), où certains $\bar{a}_{ii}$ peuvent, certes, être nuls. Compte tenu de ce qui précède, nous avons deux formules

$$\det A = (-1)^q \det \bar{A} = (-1)^q \bar{a}_{11}\bar{a}_{22} \ldots \bar{a}_{nn} \text{ (voir (6))},$$

$$\mathscr{D}(A) = (-1)^q \mathscr{D}(\bar{A}),$$

où q est le nombre de transformations élémentaires de type (I) appliquées lors du passage de A à $\bar{A}$. L'égalité $\mathscr{D}(A) = \rho \cdot \det A$, qui nous intéresse, est visiblement la conséquence de la formule

$$\mathscr{D}(\bar{A}) = \mathscr{D}(E) \cdot \bar{a}_{11} \ldots \bar{a}_{nn} \tag{7}$$

que nous allons maintenant démontrer (d'ailleurs, (6) résulte de (7), car, pour $\mathscr{D} = \det$, on aura, en raison de la propriété D3, $\mathscr{D}(E) = 1$).

Suivant la condition (i) du théorème, nous pouvons faire sortir $\bar{a}_{nn}$ du signe de $\mathscr{D}$, de sorte que

$$\mathscr{D}(\bar{A}) = \bar{a}_{nn} \mathscr{D}\left(\begin{Vmatrix} \bar{a}_{11} & \ldots & \bar{a}_{1,n-1} & \bar{a}_{1n} \\ & & & \\ 0 & \ldots & a_{n-1,n-1} & a_{n-1,n} \\ 0 & \ldots & 0 & 1 \end{Vmatrix}\right).$$

Appliquons maintenant à $\bar{A}$ une transformation élémentaire de type (II) : retranchons de la i-ième ligne de la matrice se trouvant sous le signe $\mathscr{D}$ la dernière ligne multipliée par $\bar{a}_{in}$. Il en résulte que tous les éléments de la dernière colonne s'annulent (sauf $\bar{a}_{nn} = 1$), alors que tous les autres éléments de la matrice restent inchangés. Appliquons le même raisonnement à l'avant-dernière ligne de la matrice nouvellement obtenue, et ainsi de suite. Chaque fois on fait sortir l'élément suivant $\bar{a}_{ii}$ du signe $\mathscr{D}$ et on reprend le raisonnement. En le répétant n fois, nous nous assurons que

$$\mathscr{D}(A) = \bar{a}_{nn} \ldots \bar{a}_{11} \cdot \mathscr{D}\left(\begin{Vmatrix} 1 & \ldots & 0 \\ \cdot & \cdot \; \cdot & \cdot \\ 0 & \ldots & 1 \end{Vmatrix}\right).$$

Le résultat obtenu constitue justement la formule (7). ∎

Ainsi, la fonction det se définit d'une façon unique par les propriétés D1 à D3. Voilà pourquoi, nous avons rangé ces dernières parmi les propriétés essentielles des déterminants. Dès le début, la fonction $\mathscr{D}$ vérifiant les propriétés D1 à D3 aurait pu être appelée déterminant, mais dans ce cas il aurait fallu établir son existence.

Dans notre raisonnement, cette existence est assurée par la construction même de la fonction det, c'est-à-dire par la formule (1).

Ayant en vue des applications ultérieures du théorème 2, nous n'avons pas inclus dans son énoncé la condition de norme $\mathscr{D}(E) = 1$.

EXERCICES

1. En utilisant la formule (1) et la règle des signes dans le développement d'un déterminant du troisième ordre (chap. 1, § 4, exercice 1), écrire sous forme explicite tous les produits entrant dans le développement d'un déterminant d'ordre 4. Faire attention au nombre total de termes figurant dans le développement et essayer d'établir une loi régissant la distribution des signes.

2. Le second membre de la formule (1) contient n termes. A son tour, chaque mineur M_{i1} s'écrit sous la forme d'une combinaison linéaire de ses $n-1$ mineurs d'ordre $n-2$, et ainsi de suite. Le développement d'un déterminant det (a_{ij}) d'ordre n contient donc au total $n(n-1)\ldots 3\cdot 2\cdot 1 = n!$ (factorielle n) produits de la forme $a_{i_1 1} a_{i_2 2} \ldots a_{i_n n}$ affectés de signe $+$ ou $-$. Montrer que

$$\det(a_{ij}) = a_{11}a_{22}\ldots a_{nn} + (-1)^{\frac{n(n-1)}{2}} a_{n1}a_{n-1,2}\ldots a_{1n} + \ldots$$

3. En appliquant les remarques faites dans l'exercice précédent à det (a_{ij}), avec $a_{ij} = 1$ pour $i, j = 1, 2, \ldots, n$, montrer que dans le développement de tout déterminant d'ordre n la moitié des produits $a_{i_1 1} a_{i_2 2} \ldots a_{i_n n}$ est affectée du signe $+$.

4. Ecrire la fonction symétrique gauche $\Delta: \mathbb{R}^3 \to \mathbb{R}$ de trois variables x, y, z:

$$\Delta(x, y, z) = (y-x)(z-x)(z-y)$$

sous la forme d'un déterminant du troisième ordre.

§ 2. Autres propriétés des déterminants

1. Développement suivant une colonne.— Nous sommes maintenant en mesure de répondre à la question qu'on se posait involontairement encore lors de la construction de la fonction det: la première colonne joue-t-elle un rôle particulier dans la formule récurrente (1) pour un déterminant d'ordre n? La réponse est contenue dans la formule suivante:

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} M_{ij}. \tag{1}$$

Pour sa démonstration il suffit d'appliquer le théorème 2 du § 1 à la fonction $\mathscr{D}_j$ dont il s'agit dans le lemme 2 du § 1. On obtient la relation

$$\mathscr{D}_j(A) = \mathscr{D}_j(E)\cdot\det A.$$

Or, d'après la formule (3) du § 1, on a $\mathscr{D}_j(E) = (-1)^{j-1}$. Par conséquent, $\mathscr{D}_j(A) = (-1)^{j-1} \det A$. En multipliant les deux membres de cette égalité par $(-1)^{j-1}$, on obtient $\det A = (-1)^{j-1}\mathscr{D}_j(A)$,

ce qui n'est qu'une autre écriture de la formule (1). Cette formule devient encore plus symétrique si l'on introduit un scalaire $A_{ij} = (-1)^{i+j}M_{ij}$ appelé *cofacteur* de l'élément a_{ij} dans le déterminant det A. Enonçons le résultat obtenu.

THÉORÈME 1. — *Le déterminant d'une matrice A est égal à la somme des produits de tous les éléments d'une colonne quelconque par leurs cofacteurs*

$$\det A = \sum_{i=1}^{n} a_{ij}A_{ij}. \quad \blacksquare \tag{2}$$

Dans cette proposition, toutes les colonnes jouent déjà le même rôle. Pour $j = 1$, elle se transforme en développement de départ (1) du § 1, qui introduit la notion de déterminant. On dit que les formules (1) et (2) donnent le *développement du déterminant suivant la j-ième colonne.*

On éprouve la tentation de comparer (2) avec une somme analogue $\sum_{j=1}^{n} a_{ij}A_{ij}$ suivant le deuxième indice. Nous verrons bientôt que la valeur du det A sera la même.

2. Propriétés des déterminants par rapport aux colonnes.— En considérant les diverses applications du théorème 1 nous pouvons établir toute une série de nouvelles propriétés des déterminants.

THÉORÈME 2. — *Les propriétés* D1 *à* D7 *établies au* § 1 *ont lieu non seulement par rapport aux lignes, mais aussi par rapport aux colonnes.*

DÉMONSTRATION. — Comme il a été convenu dès le début det A = det $[A_1, \ldots, A_n]$ = det $(A^{(1)}, \ldots, A^{(n)})$, et il résulte du § 1 que les propriétés D4 à D7 sont des conséquences tout à fait formelles des propriétés D1 à D3, si bien que leurs analogues pour les colonnes une fois démontrées, nous obtenons automatiquement les autres propriétés par rapport aux colonnes. Pourtant, la propriété de norme D3 occupe une position particulière et ne se rapporte ni aux lignes ni aux colonnes. Ainsi, il ne nous reste qu'à considérer les propriétés D1 et D2.

Partons de la formule (2). Elle indique directement que det A est une fonction linéaire des éléments de la j-ième colonne parce que les cofacteurs A_{ij} ne dépendent pas de ces éléments. Par là même, la propriété D1 est démontrée.

Démontrons maintenant la propriété D2 en raisonnant par récurrence sur l'ordre n du déterminant, c'est-à-dire que la fonction det $(A^{(1)}, \ldots, A^{(n)})$ est symétrique gauche. Pour $n = 1$, la propriété D2 n'a pas de sens. Pour $n = 2$ elle est facile à vérifier directe-

ment :

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc = -\begin{vmatrix} b & a \\ d & c \end{vmatrix}.$$

Soit $n > 2$. Supposons que les colonnes $A^{(k)}$ et $A^{(k+1)}$ soient permutées. Utilisons la formule (2), avec $j \neq k$, $k+1$. Le mineur M_{ij} (ou le cofacteur A_{ij}) contient les deux colonnes $A^{(k)}$, $A^{(k+1)}$, mais sous une forme raccourcie, sans éléments a_{ik}, $a_{i,k+1}$. Par hypothèse de récurrence, l'échange de deux colonnes remplace chaque mineur par son opposé. Par conséquent, on a

$$\det(\ldots, A^{(k)}, A^{(k+1)}, \ldots) = -\det(\ldots, A^{(k+1)}, A^{(k)}, \ldots). \blacksquare$$

3. Déterminant transposé. — Rappelons la notion introduite dans le chapitre 2, § 2, exercice 1. Une matrice rectangulaire à n lignes et m colonnes dont la i-ième colonne, $i = 1, 2, \ldots, m$, coïncide avec la i-ième ligne d'une matrice A à m lignes et n colonnes, s'appelle la *transposée* de la matrice A. La matrice transposée de A est notée ${}^tA$ ou A'. Donc, si $A = (a_{ij})$, ${}^tA = (a'_{ji})$, alors $a'_{ji} = a_{ij}$. Par exemple,

$${}^t\begin{Vmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{Vmatrix} = \begin{Vmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{Vmatrix}.$$

Une colonne peut être considérée comme une ligne transposée

$$[x_1, \ldots, x_n] = {}^t(x_1, \ldots, x_n).$$

Dans le cas des matrices carrés on dit aussi que le déterminant

$$\det {}^tA = \begin{vmatrix} a_{11} & a_{21} & \ldots & a_{n1} \\ a_{12} & a_{22} & \ldots & a_{n2} \\ \cdot & \cdot & \cdot & \cdot \\ a_{1n} & a_{2n} & \ldots & a_{nn} \end{vmatrix}$$

est obtenu par *transposition du déterminant* $\det A$. L'opération de transposition d'une matrice (d'un déterminant) d'ordre n peut être représentée de façon spectaculaire comme rotation de la matrice (du déterminant) autour d'un axe immobile, à savoir, autour de la diagonale principale. La rotation autour de la deuxième diagonale (non principale) est de beaucoup moins utilisée.

THÉORÈME 3. — *Le déterminant d'une matrice transposée coïncide avec celui de la matrice donnée*

$$\det {}^tA = \det A.$$

DÉMONSTRATION. — Considérons la fonction $\mathscr{D}: M_n(\mathbb{R}) \to \mathbb{R}$ obtenue par la composition $A \mapsto {}^tA \mapsto \det {}^tA$ de la fonction de passage à la matrice transposée et de la fonction det. La fonction $\mathscr{D}$ vérifie les propriétés (i), (ii) énoncées dans le théorème 2 du § 1. En effet, d'après le théorème 2, que nous venons de démontrer, la fonction ${}^tA \mapsto \det {}^tA$ possède les propriétés D1 à D7 par rapport aux colonnes de la matrice ${}^tA$, c'est-à-dire par rapport aux lignes de la matrice A. Par suite, $\mathscr{D}$ est une fonction multilinéaire symétrique gauche des lignes de la matrice. Suivant le théorème 2 du § 1, on a $\mathscr{D}(A) = \mathscr{D}(E) \cdot \det A = \det {}^tE \cdot \det A$. Mais, ${}^tE = E$ et donc $\det {}^tE = 1$, si bien que $\mathscr{D}(A) = \det A$. ∎

Conformément au théorème 3, les lignes et les colonnes d'un déterminant sont d'une même importance: les propriétés exprimées en termes de lignes s'expriment aussi en termes de colonnes et réciproquement. Par exemple, avec le théorème 1 sur le développement du déterminant suivant les éléments d'une colonne, est vrai le théorème suivant:

THÉORÈME 1′. — *Le déterminant d'une matrice A est égal à la somme des produits de tous les éléments d'une ligne fixe quelconque par leurs cofacteurs*:

$$\det A = \sum_{j=1}^{n} a_{ij}A_{ij}. \quad ∎$$

On peut y ajouter le critère suivant: *si une ligne quelconque (une colonne quelconque) du déterminant* det A *est une combinaison linéaire des autres lignes (des autres colonnes), on a* $\det A = 0$ (voir propriétés D1′, D1″ et leurs analogues pour les colonnes).

Les propriétés des déterminants que nous venons d'établir sont illustrées par deux exemples suivants:

EXEMPLE 1. — Le déterminant

$$\Delta_n = \begin{Vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ \dots & \dots & \dots & \dots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{Vmatrix} = \Delta(x_1, x_2, \dots, x_n)$$

appelé *déterminant de Vandermonde* se calcule à l'aide de la formule

$$\Delta_n = \prod_{1 \leqslant i < j \leqslant n} (x_j - x_i), \tag{3}$$

ou encore, avec une écriture plus détaillée, à l'aide de la formule

$$\Delta_n = (x_2 - x_1)(x_3 - x_1) \dots (x_n - x_1)(x_3 - x_2) \dots \\ \dots (x_n - x_2) \dots (x_n - x_{n-1})$$

(eu égard à cette formule il est utile de se reporter à l'exercice 4 du § 1).

En particulier, lorsque les éléments $x_1, \ldots, x_n$ sont deux à deux distincts, le déterminant de Vandermonde est différent de zéro. Cette propriété est bien souvent mise à profit. D'après le théorème 3, on a aussi

$$\Delta_n = \begin{vmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \ldots & x_2^{n-1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-1} \end{vmatrix}.$$

Pour démontrer la formule (3), raisonnons par récurrence sur n. En admettant que Δ_m, $m < n$, se calcule par la formule (3) et en nous appuyant sur la propriété D7, retranchons de chaque i-ième ligne du déterminant Δ_n la $(i-1)$-ième ligne multipliée par x_1:

$$\Delta_n = \begin{vmatrix} 1 & 1 & \ldots & 1 \\ 0 & x_2 - x_1 & \ldots & x_n - x_1 \\ 0 & x_2^2 - x_2 x_1 & \ldots & x_n^2 - x_n x_1 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & x_2^{n-1} - x_2^{n-2} x_1 & \ldots & x_n^{n-1} - x_n^{n-2} x_1 \end{vmatrix}.$$

Une idée vient à l'esprit de développer maintenant le déterminant Δ_n suivant la première colonne et, dans le déterminant d'ordre $n-1$ ainsi obtenu, de faire sortir du signe de déterminant le facteur commun $x_{j+1} - x_1$ de la j-ième colonne ($j = 1, 2, \ldots, n-1$) (propriété D1″ pour les colonnes). On obtient l'expression

$$\Delta_n = (x_n - x_1)(x_{n-1} - x_1) \ldots (x_2 - x_1) \begin{vmatrix} 1 & 1 & \ldots & 1 \\ x_2 & x_3 & \ldots & x_n \\ \cdot & \cdot & \cdot & \cdot \\ x_2^{n-2} & x_3^{n-2} & \ldots & x_n^{n-2} \end{vmatrix} =$$

$$= (x_n - x_1)(x_{n-1} - x_1) \ldots (x_2 - x_1) \cdot \Delta(x_2, x_3, \ldots, x_n),$$

qui coïncide avec (3), car par hypothèse de récurrence on a $\Delta(x_2, \ldots, x_n) = \prod\limits_{2 \leqslant i < j \leqslant n} (x_j - x_i)$.

EXEMPLE 2. — Une matrice $A = (a_{ij})$ de la forme

$$A = \begin{Vmatrix} 0 & a_{12} & a_{13} & \ldots & a_{1n} \\ -a_{12} & 0 & a_{23} & \ldots & a_{2n} \\ -a_{13} & -a_{23} & 0 & \ldots & a_{3n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ -a_{1n} & -a_{2n} & -a_{3n} & \ldots & 0 \end{Vmatrix}$$

est dite *symétrique gauche* ou *antisymétrique* (on dit que son déterminant est aussi *antisymétrique*). En d'autres termes, ${}^tA = -A$. Compte tenu du théorème 3, on a

$$\det A = \det {}^tA = \det(-A) = (-1)^n \det A,$$

d'où il résulte que $[1 + (-1)^{n-1}] \det A = 0$. Pour n impair on a $\det A = 0$, ce qui veut dire que le *déterminant de toute matrice antisymétrique d'ordre impair est nul*.

4. Déterminants des matrices spéciales. — Le déterminant $\det A$ d'une matrice A est d'autant plus facile à calculer que le nombre de zéros parmi les éléments de cette matrice est plus élevé et que ces

zéros sont « mieux disposés ». Cette idée intuitive trouve dans certains cas une expression quantitative précise. Nous savons par exemple (voir § 1, n° 2) que le déterminant d'une matrice triangulaire (supérieure ou inférieure) est égal au produit des éléments de sa diagonale principale. Un autre cas particulier de grande importance fait l'objet de l'assertion suivante:

THÉORÈME 4. — *Pour un déterminant D d'ordre $n+m$, dont les éléments se trouvent à l'intersection des n premières colonnes et des m dernières lignes sont zéros, est valable la formule*

$$\begin{vmatrix} a_{11} & \dots & a_{1n} & a_{1,n+1} & \dots & a_{1,n+m} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \dots & a_{nn} & a_{n,n+1} & \dots & a_{n,n+m} \\ 0 & \dots & 0 & b_{11} & \dots & b_{1m} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \dots & 0 & b_{m1} & \dots & b_{mm} \end{vmatrix} = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \cdot & \cdot & \cdot \\ a_{n1} & \dots & a_{nn} \end{vmatrix} \cdot \begin{vmatrix} b_{11} & \dots & b_{1m} \\ \cdot & \cdot & \cdot \\ b_{m1} & \dots & b_{mm} \end{vmatrix}$$

(le déterminant représenté par le premier membre de cette égalité s'appelle déterminant *quasi triangulaire* ou *déterminant à coin de zéros*).

DÉMONSTRATION. — Fixons d'abord les $n(n+m)$ éléments a_{ij} et considérons le déterminant D comme une fonction des éléments b_{kl} qui forment une matrice carrée B d'ordre m. La fonction ainsi obtenue peut être considérée comme fonction de la matrice B: $D = \mathscr{D}(B)$.

Il est clair que les propriétés du déterminant D d'être multilinéaire et symétrique gauche par rapport à m dernières lignes sont équivalentes aux mêmes propriétés de $\mathscr{D}(B)$ par rapport aux lignes de la matrice B. Cela signifie que nous sommes en droit d'appliquer à $\mathscr{D}(B)$ le théorème 2 du § 1 d'après lequel $\mathscr{D}(B) = \mathscr{D}(E)\cdot\det B$. Par la définition de la fonction $\mathscr{D}$ on a

$$\mathscr{D}(E) = \begin{vmatrix} a_{11} & \dots & a_{1n} & a_{1,n+1} & \dots & a_{1,n+m} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \dots & a_{nn} & a_{n,n+1} & \dots & a_{n,n+m} \\ 0 & \dots & 0 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \dots & 0 & 0 & \dots & 1 \end{vmatrix}.$$

Développons $\mathscr{D}(E)$ suivant les éléments de la dernière ligne (voir formule (2)), puis suivant les éléments de l'avant-dernière ligne, et ainsi de suite. En réitérant cette opération m fois, nous nous assu-

rons que $\mathscr{D}(E) = \det A$, où

$$A = \begin{Vmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{Vmatrix}.$$

Finalement, on obtient $D = \mathscr{D}(B) = \det A \cdot \det B$. ■

Avec les nouvelles notations, la formule du théorème 4 prend une forme plus compacte

$$\det \begin{Vmatrix} A & C \\ 0 & B \end{Vmatrix} = \det A \cdot \det B. \tag{4}$$

Ici A et B sont des matrices carrées, alors que la matrice 0 et la matrice C sont rectangulaires. En s'appuyant sur les théorèmes 3 et 4 ou sur le raisonnement développé au cours de la démonstration du théorème 4, on établit sans peine que

$$\det \begin{Vmatrix} A & 0 \\ C & B \end{Vmatrix} = \det A \cdot \det B. \ ■$$

Parfois on tente d'écrire exactement la même expression pour le déterminant $\det \begin{Vmatrix} C & A \\ B & 0 \end{Vmatrix}$, bien qu'un contre-exemple très simple $\begin{Vmatrix} 0 & 1 \\ 1 & 0 \end{Vmatrix} = -1$ vienne naturellement à l'esprit. Tout dépend du signe. On obtient la réponse correcte à condition de faire l'échange des lignes ou des colonnes qui ramène la matrice $\begin{Vmatrix} C & A \\ B & 0 \end{Vmatrix}$ à la forme $\begin{Vmatrix} B & 0 \\ C & A \end{Vmatrix}$ ou $\begin{Vmatrix} A & C \\ 0 & B \end{Vmatrix}$.

On peut développer des raisonnements plus simples en partant du même théorème 2 du § 1 que nous avons déjà utilisé à maintes reprises. En effet,

$$\det \begin{Vmatrix} C & A \\ B & 0 \end{Vmatrix} = \det \begin{Vmatrix} C & A \\ E & 0 \end{Vmatrix} \cdot \det B.$$

D'après la formule (1) appliquée m fois, on trouve

$$\det \begin{Vmatrix} C & A \\ E & 0 \end{Vmatrix} = \begin{vmatrix} & * & & a_{11} & \dots & a_{1n} \\ & & & a_{n1} & \dots & a_{nn} \\ 1 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & 0 & & 0 \end{vmatrix} =$$

$$= (-1)^{(n+2)+(n+2)+\dots+(n+2)} \det A = (-1)^{nm} \det A.$$

Il s'ensuit définitivement que si A, B sont des matrices carrées d'ordres respectifs n et m, on a

$$\det \begin{Vmatrix} C & A \\ B & 0 \end{Vmatrix} = (-1)^{nm} \det A \cdot \det B. \tag{5}$$

Les formules (4) et (5) font partie du théorème général de Laplace sur le développement des déterminants. Pourtant, ce théorème n'étant utilisé que dans des cas relativement rares, nous ne nous y étendrons pas. Nous ne nous

empressons pas non plus de démontrer le théorème dit théorème sur le développement complet du déterminant (voir chap. 4, § 3) qui n'est, au point de vue du calcul, que de peu d'utilité.

Une assertion très importante, relative aux déterminants des matrices, est contenue dans le théorème suivant :

THÉORÈME 5. — *Soient A et B deux matrices carrées d'ordre n. Alors, on a*

$$\det AB = \det A \cdot \det B.$$

DÉMONSTRATION. — Suivant les formules (7) et (9) du chapitre 2, § 3, qui expriment les coefficients c_{ij} de la matrice $(c_{ij}) = AB = (a_{ij})(b_{ij})$ par les coefficients des matrices A et B, la i-ième ligne $(AB)_i$ s'écrit sous la forme

$$(AB)_i = (A_iB^{(1)}, A_iB^{(2)}, \ldots, A_iB^{(n)}); \quad A_iB^{(j)} = \sum_{k=1}^{n} a_{ik}b_{kj}.$$

Fixons la matrice B et posons pour toute matrice A :

$$\mathscr{D}(A) = \det AB.$$

Montrons que la fonction $\mathscr{D}$ satisfait aux conditions (i), (ii) du théorème 2 du § 1. En effet, nous savons que $\det AB$ est une fonction linéaire des éléments de la i-ième ligne $(AB)_i$:

$$\det AB = \lambda_1 A_iB^{(1)} + \lambda_2 A_iB^{(2)} + \ldots + \lambda_n A_iB^{(n)}.$$

Par suite,

$$\mathscr{D}(A) = \sum_{j=1}^{n} \lambda_j \sum_{k=1}^{n} a_{ik}b_{kj} = \sum_{k=1}^{n} a_{ik} \sum_{j=1}^{n} \lambda_j b_{kj} = \sum_{k=1}^{n} \mu_k a_{ik},$$

où $\mu_k = \sum_{j=1}^{n} \lambda_j b_{kj}$ est un scalaire qui ne dépend pas des éléments de la i-ième ligne A_i de la matrice A.

Nous voyons que $\mathscr{D}(A)$ dépend linéairement des éléments de la i-ième ligne de la matrice A.

Permutons A_s et A_t. Puisque la s-ième et la t-ième lignes de la matrice AB sont de la forme

$$(A_sB^{(1)}, \ldots, A_sB^{(n)}),$$
$$(A_tB^{(1)}, \ldots, A_tB^{(n)}),$$

elles seront interverties, elles aussi, si bien que, suivant le théorème 1, on aura :

$$\mathscr{D}(\ldots, A_s, \ldots, A_t, \ldots) = \mathscr{D}(A) = \det AB =$$
$$= \det [\ldots, (AB)_s, \ldots, (AB)_t, \ldots] =$$
$$= -\det [\ldots, (AB)_t, \ldots, (AB)_s, \ldots] =$$
$$= -\mathscr{D}(\ldots, A_t, \ldots, A_s, \ldots).$$

Ainsi, les deux conditions du théorème 2 du § 1, d'après lequel $\mathcal{D}(A) = \mathcal{D}(E) \cdot \det A$, sont satisfaites. Or, par définition, $\mathcal{D}(E) = \det EB = \det B$, ce qui implique la formule cherchée. ∎

5. Sur la construction de la théorie des déterminants.— Les théorèmes 1 et 2 du § 1 donnent au fond une description axiomatique de la fonction det, bien que nous ayons commencé par donner cette fonction d'une manière purement constructive.

Indiquons encore une voie à suivre pour élaborer la théorie des déterminants. A savoir, soit donnée une fonction $\mathcal{D} : M_n(\mathbb{R}) \to \mathbb{R}$ qui vérifie les propriétés suivantes:

(i) $\mathcal{D}(AB) = \mathcal{D}(A) \cdot \mathcal{D}(B)$ *quelles que soient les matrices* $A, B \in M_n(\mathbb{R})$;

(ii) $\mathcal{D}(F_{s,t}) = -1$ *pour toute matrice élémentaire* $F_{s,t}$ (voir chap. 2, § 4, nº 4);

(iii) $\mathcal{D}(A) = \lambda$ *pour une matrice triangulaire supérieure de la forme*

$$A = \begin{vmatrix} \lambda & & & * \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{vmatrix}, \quad \lambda \in \mathbb{R}.$$

On affirme que $\mathcal{D} = \det$. En effet, en utilisant la propriété (i) relativement à la matrice

$$F_s(\lambda) = F_{1,s} \begin{Vmatrix} \lambda & & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{Vmatrix} F_{1,s} = \begin{Vmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & \lambda & & \\ & & & \ddots & \\ 0 & & & & 1 \end{Vmatrix},$$

nous obtenons $\mathcal{D}(F_s(\lambda)) = (-1) \cdot \lambda \cdot (-1) = \lambda$. Cela signifie en particulier que $\mathcal{D}(F_s(\lambda)) = \lambda$ pour la matrice élémentaire $F_s(\lambda)$, $\lambda \neq 0$. On a ensuite

$$\begin{Vmatrix} E_r & 0 \\ 0 & 0 \end{Vmatrix} = F_{r+1}(0) \ldots F_n(0),$$

d'où

$$\mathcal{D}\left(\begin{Vmatrix} E_r & 0 \\ 0 & 0 \end{Vmatrix}\right) = \begin{cases} 0 \text{ pour } r < n, \\ 1 \text{ pour } r = n. \end{cases}$$

Suivant (iii), on a $\mathcal{D}(F_{s,t}(\lambda)) = 1$ pour la matrice élémentaire $F_{s,t}(\lambda)$, avec $s < t$. Puisque

$$F_{s,t} \cdot F_{s,t}(\lambda) \cdot F_{s,t} = F_{t,s}(\lambda),$$

on a $\mathcal{D}(F_{t,s}(\lambda)) = 1$, donc $\mathcal{D}(F_{s,t}(\lambda)) = 1$, quels que soient les indices $s \neq t$.

Ainsi, $\mathcal{D}(F_{s,t}) = -1 = \det F_{s,t}$, $\mathcal{D}(F_{s,t}(\lambda)) = 1 = \det F_{s,t}(\lambda)$ et $\mathcal{D}(F_s(\lambda)) = \lambda = \det F_s(\lambda)$. Puisque toute matrice $A \in M_n(\mathbb{R})$ s'écrit sous la forme $A = P \begin{Vmatrix} E_r & 0 \\ 0 & 0 \end{Vmatrix} Q$, $r \leqslant n$, où P et Q sont des produits de matrices élémentaires (voir les raisonnements qui précèdent le théorème 5 du chap. 2, § 4), la propriété (i) permet de conclure que $\mathcal{D}(A) = \det A$.

Nous conseillons au lecteur d'avancer et de justifier ses propres variantes de description axiomatique de la fonction det.

EXERCICES

1. Les entiers $1798 = 31 \cdot 58$, $2139 = 31 \cdot 69$, $3255 = 31 \cdot 105$, $4867 = 31 \cdot 157$ sont divisibles par 31. Sans procéder à aucun calcul, montrer que le déterminant d'ordre 4

$$\begin{vmatrix} 1 & 7 & 9 & 8 \\ 2 & 1 & 3 & 9 \\ 3 & 2 & 5 & 5 \\ 4 & 8 & 6 & 7 \end{vmatrix}$$

est, lui aussi, divisible par 31.

2. Montrer que tout déterminant antisymétrique $|a_{ij}|$ d'ordre 4, avec $a_{ij} \in \mathbb{Z}$, est le carré d'un entier. (Remarque. Cela est vrai pour tout déterminant antisymétrique.)

3. Démontrer la relation $\det AB = \det A \cdot \det B$ (théorème 5) en ramenant une matrice auxiliaire $C = \begin{Vmatrix} E & B \\ -A & 0 \end{Vmatrix}$ de type $(2n, 2n)$ à la forme $C' = \begin{Vmatrix} E & B \\ 0 & AB \end{Vmatrix}$ par application à ses lignes des transformations élémentaires de type (II). (Indication. Utiliser l'égalité $\det C = \det C'$ et les relations (4), (5).)

4. Montrer que ${}^t(AB) = {}^tB\,{}^tA$ quelles que soient les matrices rectangulaires $m \times r$ et $r \times n$.

5. Montrer que $\det B^{-1}AB = \det A$ pour toute matrice carrée $A \in M_n(\mathbb{R})$ et toute matrice inversible $B \in M_n(\mathbb{R})$.

6. Soit

$$C_n(\lambda_1, \ldots, \lambda_n) = \begin{Vmatrix} \lambda_1 & 1 & 0 & \ldots & 0 & 0 & 0 \\ -1 & \lambda_2 & 1 & \ldots & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & \lambda_{n-2} & 1 & 0 \\ 0 & 0 & 0 & \ldots & -1 & \lambda_{n-1} & 1 \\ 0 & 0 & 0 & \ldots & 0 & -1 & \lambda_n \end{Vmatrix}.$$

Montrer que $\det C_n = \lambda_n \det C_{n-1} + \det C_{n-2}$. Calculer la valeur numérique de $\det C_n$ pour $\lambda_1 = \lambda_2 = \ldots = \lambda_n = 1$. (Indication. Se rappeler de l'exemple 3 du chap. 2, § 3, nº 3, et faire attention au fait que $\det C_n(1, \ldots, 1) = (-1)^{n-1} \det C_n(-1, \ldots, -1)$.)

7. Montrer que le déterminant de la matrice $n \times n$

$$A_n = \begin{Vmatrix} 2 & -1 & 0 & 0 & \ldots & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & \ldots & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & \ldots & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \ldots & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & \ldots & 0 & -1 & 2 \end{Vmatrix}$$

est égal à $n + 1$.

§ 3. Applications des déterminants

1. Critère de régularité d'une matrice.— Comme il a été dit au chapitre 2, § 3, une matrice carrée A est régulière s'il existe une matrice inverse A^{-1}. En appliquant le théorème 5 du § 2 à la rela-

tion $AA^{-1} = A^{-1}A = E$, on obtient $\det A \cdot \det A^{-1} = 1$. Il s'ensuit que le *déterminant d'une matrice régulière est différent de zéro et*

$$\det A^{-1} = (\det A)^{-1}.$$

Considérons avec la matrice A son *adjointe*

$$A^{\vee} = \left\| \begin{matrix} A_{11} & \dots & A_{n1} \\ \dots & \dots & \dots \\ A_{1n} & \dots & A_{nn} \end{matrix} \right\|.$$

Pour obtenir $A^{\vee}$ il faut mettre à la place de chaque élément a_{ij} de la matrice A son cofacteur A_{ij} $(i, j = 1, \dots, n)$ et passer ensuite à la matrice transposée.

THÉORÈME 1. — *Une matrice* $A \in M_n(\mathbb{R})$ *est régulière (inversible) si, et seulement si,* $\det A \neq 0$. *Si* $\det A \neq 0$, *on a* $A^{-1} = (\det A)^{-1}A^{\vee}$ *ou bien, sous une forme plus détaillée,*

$$\left\| \begin{matrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{matrix} \right\|^{-1} = \left\| \begin{matrix} \frac{A_{11}}{\det A} & \dots & \frac{A_{n1}}{\det A} \\ \dots & \dots & \dots \\ \frac{A_{1n}}{\det A} & \dots & \frac{A_{nn}}{\det A} \end{matrix} \right\|.$$

La démonstration de ce théorème sera précédée du lemme suivant :

LEMME. — *Soit* $A \in M_n(\mathbb{R})$. *Alors, on a les relations*

$$a_{i1}A_{j1} + a_{i2}A_{j2} + \dots + a_{in}A_{jn} = \delta_{ij} \det A, \tag{1}$$

$$a_{1i}A_{1j} + a_{2i}A_{2j} + \dots + a_{ni}A_{nj} = \delta_{ij} \det A, \tag{2}$$

où δ_{ij} *est le symbole de Kronecker (pour* $i \neq j$ *on dit que le déterminant est développé suivant les éléments d'une ligne non propre ou respectivement d'une colonne non propre).*

DÉMONSTRATION. — Pour $i = j$, l'assertion énoncée dans le lemme coïncide avec les théorèmes 1 et 1′ du § 2. Par suite, il reste à considérer le cas de $i \neq j$, où $\delta_{ij} = 0$. A cet effet, introduisons une matrice

$$A' = [A_1, \dots, A_i, \dots, A_i, \dots, A_n] = \left\| \begin{matrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{matrix} \right\|,$$

obtenue à partir de $A = [\ldots, A_i, \ldots, A_j, \ldots]$ en remplaçant la j-ième ligne par la i-ième (la i-ième ligne reste sur place). Cette matrice, comme toute matrice carrée contenant deux lignes identiques, a $\det A' = 0$. D'autre part, le cofacteur A'_{jk} ($k = 1, \ldots, n$) est obtenu en supprimant la j-ième ligne $A'_j = A_i$ et la k-ième colonne du déterminant, de sorte que $A'_{jk} = A_{jk}$. Le développement formel du déterminant de la matrice $A' = (a'_{st})$ suivant la j-ième ligne donne la relation

$$0 = \det A' = \sum_{k=1}^{n} a'_{jk}A'_{jk} = \sum_{k=1}^{n} a_{ik}A_{jk},$$

qui coïncide avec la relation (1) dans l'énoncé du lemme. La deuxième relation est obtenue à partir des considérations analogues relatives aux colonnes. ■

En revenant à la démonstration du théorème, nous remarquons simplement que le premier membre de la relation (1) n'est rien d'autre que l'élément c_{ij} de la matrice $C = AA^\vee$:

$$\left\|\begin{matrix} c_{11} & \ldots & c_{1n} \\ \ldots & \ldots & \ldots \\ c_{n1} & \ldots & c_{nn} \end{matrix}\right\| = \left\|\begin{matrix} a_{11} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{nn} \end{matrix}\right\| \left\|\begin{matrix} A_{11} & \ldots & A_{n1} \\ \ldots & \ldots & \ldots \\ A_{1n} & \ldots & A_{nn} \end{matrix}\right\|.$$

D'après la relation (1) on a $(c_{ij}) = (\delta_{ij} \det A) = (\det A)\, E$. Ainsi

$$AA^\vee = (\det A)\, E,$$

d'où l'on déduit pour $\det A \neq 0$:

$$(\det A)^{-1} (AA^\vee) = A\, (\det A)^{-1}A^\vee = E.$$

Le premier membre de la relation (2) est l'expression de l'élément c'_{ji} de la matrice $C' = A^\vee A$. Puisque les seconds membres des relations (1) et (2) coïncident, nous obtenons, dans le cas où $\det A \neq$ $\neq 0$, les relations

$$A\, (\det A)^{-1} A^\vee = (\det A)^{-1}A^\vee A = E,$$

qui signifient que $A^{-1} = (\det A)^{-1}A^\vee$. ■

Corollaire 1. — *Le déterminant est nul si, et seulement si, ses lignes (et ses colonnes) sont linéairement dépendantes.*

Ce critère, que nous connaissons déjà en partie (voir la fin du nº 3, § 2), aurait pu être démontré depuis longtemps, mais nous n'en avions pas besoin. En voici la démonstration : suivant le théorème 1, l'égalité $\det A = 0$ est équivalente à la singularité de la matrice A, alors que d'après le théorème 4 du chapitre 2, § 3, la singularité est équivalente à la condition rang $A < n$ (n est l'ordre de la matrice carrée A) qui caractérise, en vertu du théorème 1 du chapitre 2, § 2, les matrices à lignes (colonnes) linéairement dépendantes. ■

Le théorème 1 est d'une portée plutôt théorique. Au point de vue du calcul, surtout dans le cas des matrices d'ordre élevé, le calcul de la matrice A^{-1} s'avère plus commode en appliquant la méthode de la (P, Q)-réduction, décrite dans le corollaire au théorème 5 du chapitre 2, § 4.

Proposons-nous maintenant d'établir les formules pour la résolution d'un système linéaire de n équations à n inconnues ; d'ailleurs, initialement la théorie des déterminants a été développée justement pour la résolution de tels systèmes.

COROLLAIRE 2 (formules de Cramer). — *Si un système linéaire*

$$\begin{aligned} a_{11}x_1 + \ldots + a_{1n}x_n &= b_1, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + \ldots + a_{nn}x_n &= b_n \end{aligned}$$

possède un déterminant non nul (c'est-à-dire $\det(a_{ij}) \neq 0$*) il admet une solution et une seule, définie par la formule*

$$x_k^\circ = \frac{\begin{vmatrix} a_{11} & \ldots & b_1 & \ldots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \ldots & b_n & \ldots & a_{nn} \end{vmatrix}}{\begin{vmatrix} a_{11} & \ldots & a_{1k} & \ldots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \ldots & a_{nk} & \ldots & a_{nn} \end{vmatrix}}, \quad k = 1, 2, \ldots, n$$

(le numérateur D_k est obtenu en remplaçant dans $D = \det(a_{ij})$ la k-ième colonne par la colonne des termes constants).

DÉMONSTRATION. — Suivant le théorème 1, la matrice $A = (a_{ij})$ est inversible. Par conséquent, en écrivant notre système sous la forme $AX = B$, nous obtenons, de même qu'au chapitre 2, § 4 :

$$X = A^{-1}B = \frac{(A_{ji})\,B}{\det A}.$$

Il en résulte

$$x_k^\circ = \frac{\sum_j A_{jk}b_j}{\det A},$$

où l'expression au numérateur représente tout juste le développement du déterminant D_k suivant la k-ième colonne (voir (2)).

L'exécution de toutes les transformations en ordre inverse montre que l'ensemble $(D_1/\det A, \ldots, D_n/\det A)$ est réellement solution de notre système. ■

Remarquons que les formules (3), (9) du chapitre 1, § 4, coïncident justement avec les formules de Cramer pour $n = 2$ et 3 respectivement. Les formules de Cramer, qui s'avèrent bien commodes pour de faibles valeurs de n, ne remplissent au fond qu'une fonction purement théorique. C'est ainsi, par exemple, que

leur application au système linéaire indiqué dans l'exemple 2 (chap. 1, § 3, n° 4) donne pour les nombres de Fibonacci (en tenant compte de l'égalité det $A = 1$) l'expression suivante:

$$f_n = \begin{vmatrix} 1 & 0 & 0 & \dots & 0 & 0 & 1 \\ 0 & 1 & 0 & \dots & 0 & 0 & 1 \\ -1 & -1 & 1 & \dots & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & -1 & 1 & 0 \\ 0 & 0 & 0 & \dots & -1 & -1 & 0 \end{vmatrix}.$$

On comprend qu'elle est assez loin de l'expression explicite obtenue pour f_n à la fin du § 3, chapitre 2.

2. Détermination du rang d'une matrice.— Les §§ 2 et 4 du chapitre 2 contiennent tout le nécessaire pour décrire l'ensemble des solutions d'un système rectangulaire général d'équations linéaires. Le rôle fondamental dans cette description appartient à la notion de rang d'une matrice. Il ne nous reste qu'à la traduire en langage de la théorie des déterminants pour avoir à notre disposition encore une méthode pour le calcul du rang et un moyen commode pour exprimer le fait d'indépendance linéaire d'un système de vecteurs de l'espace vectoriel $\mathbb{R}^m$.

Ainsi, soit

$$A = \begin{Vmatrix} a_{11} \dots a_{1r} \dots a_{1n} \\ \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \\ a_{r1} \dots a_{rr} \dots a_{rn} \\ \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \\ a_{m1} \dots a_{mr} \dots a_{mn} \end{Vmatrix}$$

une matrice rectangulaire quelconque $m \times n$ à coefficients $a_{ij} \in \mathbb{R}$. Par mineur d'ordre k de la matrice A on entend, comme à l'ordinaire, le déterminant de la matrice extraite de A, dont les éléments se trouvent à l'intersection des k colonnes différentes et des k lignes différentes marquées; $k \leqslant \min(n, m)$.

Supposons que le rang de la matrice A soit égal à r. En vertu du théorème 1 (voir chap. 2, § 2) cela signifie que r est le nombre, le plus grand possible, de lignes linéairement indépendantes, ainsi que celui de colonnes linéairement indépendantes de la matrice A. Se reportant au théorème 5 du chapitre 2, § 4, et à son corollaire, on remarque que

$$A = B \begin{Vmatrix} E_r & 0 \\ 0 & 0 \end{Vmatrix} C,$$

où B et C sont des matrices régulières $m \times m$ et $n \times n$ écrites sous la forme du produit de matrices élémentaires. Du fait que la matrice $\begin{Vmatrix} E_r & 0 \\ 0 & 0 \end{Vmatrix}$ a un mineur non nul $M = |E_r| = 1$ d'ordre r, mais n'a

pas de mineurs non nuls d'ordre $> r$, et vu que cette propriété se conserve lors des transformations élémentaires faites sur les lignes et les colonnes, nous pouvons énoncer l'assertion suivante :

THÉORÈME 2. — *Le rang de toute matrice A à m lignes et n colonnes est égal au plus grand ordre de ses mineurs non nuls.* ■

Tout mineur non nul d'ordre *maximal* de la matrice A s'appelle son *mineur de base*. Les colonnes (respectivement les lignes) de la matrice A qui croisent le mineur de base donné s'appellent, conformément à la terminologie introduite au chapitre 2, colonnes de base (respectivement, lignes de base). En interprétant, comme précédemment, les lignes et les colonnes de la matrice A de type (m, n) en tant que vecteurs des espaces respectifs $\mathbb{R}^n$ et $\mathbb{R}^m$, ainsi qu'en utilisant les propriétés fondamentales des systèmes de vecteurs linéairement indépendants (possibilité de les compléter jusqu'à la base; voir chap. 2, § 1, exercice 5), il est facile de comprendre que la recherche d'au moins un mineur de base se simplifie considérablement, lorsqu'on la ramène à une étude successive des *mineurs* dits *bordants*. A savoir, si l'on trouve un mineur non nul M d'ordre k de la matrice A, le pas suivant consiste à vérifier les seuls mineurs d'ordre $(k + 1)$, dont le mineur M est obtenu en supprimant une ligne et une colonne. Si les mineurs bordant M sont nuls, on a rang $A = k$. (Pourquoi ? Cela signifierait, en vertu du théorème 2, que toute colonne de la matrice A s'exprime linéairement en fonction des k colonnes choisies.) Dans le cas contraire, il convient de passer aux mineurs bordant un mineur non nul quelconque d'ordre $k + 1$.

La méthode des mineurs bordants s'avère assez commode en pratique, surtout dans les cas où l'on veut déterminer non seulement le rang, mais aussi les colonnes et les lignes de la matrice A qui forment un système libre maximal. Lors des transformations élémentaires, cette information est évidemment perdue.

EXERCICES

1. Montrer que les relations suivantes sont vraies :

$$(AB)^\vee = B^\vee A^\vee; \quad ({}^tA)^\vee = {}^t(A^\vee); \quad (\lambda A)^\vee = \lambda^{n-1}A^\vee;$$

$$(A^\vee)^\vee = (\det A)^{n-2}A.$$

2. Exprimer rang $A^\vee$ en fonction de rang A.
3. Démontrer qu'un système carré d'équations linéaires homogènes admet des solutions non triviales si, et seulement si, le déterminant du système est nul.
4. En s'appuyant sur les résultats obtenus au chapitre 2, § 4, n° 1, et sur le corollaire 2 du théorème 1, montrer que le système fondamental de solutions

du système homogène

$$a_{11}x_1 + \ldots + a_{1n}x_n = 0,$$
$$\ldots\ldots\ldots\ldots\ldots$$
$$a_{n-1,\,1}x_1 + \ldots + a_{n-1,\,n}x_n = 0$$

de rang $r = n - 1$ ne comprend qu'un seul vecteur colonne

$$X^\circ = [D_1, -D_2, D_3, \ldots, (-1)^{n-1} D_n],$$

où D_i est le déterminant d'une matrice obtenue à partir de $A = (a_{ij})$ en supprimant sa i-ième colonne. Toute solution du système est de la forme $X = \lambda X^\circ$.

5. Soient $A = (a_{ij}) \in M_n(\mathbb{R})$ et $(n-1) \mid a_{ij} \mid < \mid a_{ii} \mid$ pour tout $i \neq j$. Démontrer que $\det A \neq 0$. (Indication. En raisonnant par l'absurde, utiliser le critère établi dans l'exercice 3. A savoir, si $[x_1^\circ, \ldots, x_n^\circ]$ est une solution non triviale du système linéaire $AX = 0$ et x_k° est sa composante, possédant un module maximal, alors de la k-ième équation $a_{kk}x_k^\circ + \sum_{j\neq k} a_{kj}x_j^\circ = 0$ il résulte l'estimation $(n-1) \mid a_{kk} \mid \mid x_k^\circ \mid = (n-1) \mid \sum_{j\neq k} a_{kj}x_j^\circ \mid <$ $< (n-1) \mid a_{kk} \mid \mid x_k^\circ \mid$, qui fournit la contradiction requise.)

6. Démontrer l'assertion suivante. Soient $A = (a_{ij})$, $B = (b_{kl})$ deux matrices respectives $n \times m$ et $m \times n$ et soit $C = AB$. Alors

$$\det C = \sum_{1 \leqslant j_1 < \ldots < j_n \leqslant m} \begin{vmatrix} a_{1j_1} & a_{2j_1} & \ldots & a_{nj_1} \\ a_{1j_2} & a_{2j_2} & \ldots & a_{nj_2} \\ \ldots & \ldots & \ldots & \ldots \\ a_{1j_n} & a_{2j_n} & \ldots & a_{nj_n} \end{vmatrix} \begin{vmatrix} b_{j_1 1} & b_{j_1 2} & \ldots & b_{j_1 n} \\ b_{j_2 1} & b_{j_2 2} & \ldots & b_{j_2 n} \\ \ldots & \ldots & \ldots & \ldots \\ b_{j_n 1} & b_{j_n 2} & \ldots & b_{j_n n} \end{vmatrix}.$$

La somme dans le second membre est étendue à toutes les $\binom{m}{n}$ combinaisons possibles $\{j_1, j_2, \ldots, j_n\}$ de m nombres: $1, 2, \ldots, m$, pris n à n. En particulier, $\det C = \det A \cdot \det B$ pour $m = n$, et $\det C = 0$ pour $n > m$. (Indication. Répétée plusieurs fois, l'application de la propriété $D1''$ (§ 1, n° 2) donne par suite de $C = (c_{ij})$, $c_{ij} = \sum a_{ik}b_{kj}$:

$$\det C = \sum_{k_1, \ldots, k_n = 1}^{m} \begin{vmatrix} a_{1k_1} & a_{1k_2} & \ldots & a_{1k_n} \\ a_{2k_1} & a_{2k_2} & \ldots & a_{2k_n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{nk_1} & a_{nk_2} & \ldots & a_{nk_n} \end{vmatrix} b_{k_1 1} b_{k_2 2} \ldots b_{k_n n},$$

où la somme est étendue à tous les $k_1, \ldots, k_n$ deux à deux distincts. Pour $m < n$, il n'y a pas de tels indices et donc $\det C = 0$. Si $m \geqslant n$, alors $k_1, \ldots$ $\ldots, k_n$ est une suite des éléments de $\{j_1, \ldots, j_n\}$, pris dans un ordre quelconque et choisis parmi $1, 2, \ldots, m$. Il convient de réunir tous les termes correspondant à la combinaison fixée $\{j_1, \ldots, j_n\}$ et, en se servant du théorème

sur le développement complet du déterminant, obtenir l'expression

$$\sum \begin{vmatrix} a_{1k_1} & \cdots & a_{nk_1} \\ \cdot & \cdot & \cdot \\ a_{1k_n} & \cdots & a_{nk_n} \end{vmatrix} b_{k_1 1} \cdots b_{k_n n} = \begin{vmatrix} a_{1j_1} & \cdots & a_{nj_1} \\ \cdot & \cdot & \cdot \\ a_{1j_n} & \cdots & a_{nj_n} \end{vmatrix} \cdot \sum_{\pi} \varepsilon_{\pi} b_{k_1 1} \cdots b_{k_n n} =$$

$$= \begin{vmatrix} a_{1j_1} & \cdots & a_{nj_1} \\ \cdot & \cdot & \cdot \\ a_{1j_n} & \cdots & a_{nj_n} \end{vmatrix} \cdot \begin{vmatrix} b_{j_1 1} & \cdots & b_{j_1 n} \\ \cdot & \cdot & \cdot \\ b_{j_n 1} & \cdots & b_{j_n n} \end{vmatrix}, \text{ où } \pi = \begin{pmatrix} j_1 \cdots j_n \\ k_1 \ldots k_n \end{pmatrix}. \ \rangle$$

7. En utilisant l'exercice précédent, montrer que si M est une matrice $m \times n$ à coefficients dans $\mathbb{R}$, $m \geqslant n$, on a

$$\det {}^tAA = \sum_{M} M^2,$$

où M parcourt tous les $\binom{m}{n}$ mineurs d'ordre n de la matrice A.

CHAPITRE 4

STRUCTURES ALGÉBRIQUES
(groupes, anneaux, corps)

Dans les chapitres qui précèdent nous avons accumulé pas mal de données concrètes qu'il s'agit maintenant d'interpréter en nous plaçant à un point de vue plus abstrait. A cet effet, nous introduirons, en attendant à un niveau élémentaire, les notions de groupe, d'anneau et de corps qui sont des notions fondamentales pour toute l'algèbre.

§ 1. Ensembles munis d'opérations algébriques

1. Opérations binaires.— Soit X un ensemble quelconque. On appelle *opération algébrique binaire* (ou *loi de composition*) définie sur X une application arbitraire (fixe) $\tau: X \times X \to X$ du carré cartésien $X^2 = X \times X$ dans X. Ainsi à tout couple (a, b) d'éléments $a, b \in X$, cette opération fait correspondre univoquement un élément déterminé $\tau(a, b)$ du même ensemble X. Quelquefois, au lieu de $\tau(a, b)$ on écrit $a\tau b$, mais le plus souvent on désigne l'opération binaire sur X par un symbole spécial quelconque: $*$, $\circ$, $\cdot$ ou $+$. Nous suivrons, nous aussi, cette dernière voie en appelant $a \cdot b$ (ou tout simplement ab sans aucun signe entre a et b) le *produit* et $a + b$ la *somme des éléments* $a, b \in X$. On comprend que dans la plupart des cas ces dénominations sont conventionnelles.

En général, beaucoup de différentes opérations peuvent être définies sur un ensemble X. En voulant dégager l'une d'elles, on utilise les parenthèses $(X, *)$ et on dit que l'opération $*$ confère à X une *structure algébrique* ou encore que $(X, *)$ est un *système algébrique*. C'est ainsi par exemple que sur l'ensemble $\mathbb{Z}$ des entiers relatifs, en plus des opérations naturelles $+$, $\cdot$ (addition et multiplication) il est aisé d'indiquer des opérations « dérivées »: $n \circ m = n + m - nm$, $n * m = -n - m$, etc., qu'on obtient aisément à l'aide de $+$ (ou $-$) et $\cdot$. Il en résulte des structures algébriques différentes $(\mathbb{Z}, +)$, $(\mathbb{Z}, \cdot)$, $(\mathbb{Z}, \circ)$, $(\mathbb{Z}, *)$.

A côté des opérations algébriques binaires, on a aussi intérêt à considérer des opérations n-aires beaucoup plus générales (unaires pour $n = 1$, ternaires pour $n = 3$, etc.), de même que leurs combi-

naisons. Les structures algébriques associées à de telles opérations font l'objet de la théorie des algèbres universelles. D'ailleurs, nous ne les mentionnons que pour souligner une fois de plus l'importance de principe que présentent pour les mathématiques les structures algébriques munies d'opérations binaires, qui ne constituent au fond qu'une branche de la théorie des algèbres universelles.

Quant aux modes de formation des diverses opérations binaires sur un ensemble X, ils donnent évidemment libre cours à l'imagination. Mais le problème d'étude des structures algébriques arbitraires est trop général pour qu'il puisse présenter une importance réelle. C'est la raison pour laquelle on ne le considère qu'en imposant diverses limitations naturelles.

2. Demi-groupes et monoïdes.— Une opération binaire $*$ sur un ensemble X est dite *associative* si $(a * b) * c = a* (b * c)$ pour tout triplet $a, b, c \in X$; elle est *commutative* si $a * b = b * a$. On donne les mêmes noms à la structure algébrique $(X, *)$ définie par cette opération. Les conditions d'associativité et de commutativité sont indépendantes l'une de l'autre. En effet, l'opération $*$ définie sur $\mathbb{Z}$ par la loi $n * m = -n - m$ est évidemment commutative, mais $(1 * 2) * 3 = (-1-2) * 3 = -(-1-2) - 3 = 0 \neq 4 = 1 * (2 * 3)$ de sorte que la condition d'associativité n'est pas satisfaite. L'opération de multiplication définie sur l'ensemble $M_n(\mathbb{R})$ de toutes les matrices carrées d'ordre $n > 1$ est associative sans être commutative (voir chap. 2, § 3, n° 2).

Un élément $e \in X$ est dit *élément neutre* (ou *élément unité*) pour l'opération binaire considérée $*$ si $e * x = x * e = x$ pour tout $x \in X$. Si e' est un autre élément neutre, alors, en vertu de la définition, on a $e' = e' * e = e$. Cela signifie que si une structure algébrique donnée $(X, *)$ admet un élément neutre cet élément est unique.

Un ensemble X muni d'une opération binaire associative s'appelle *demi-groupe*. Un demi-groupe ayant un élément neutre (élément unité) s'appelle encore par convention *monoïde* (ou tout simplement *demi-groupe unitaire*).

De même que pour tout ensemble, la puissance d'un monoïde $M = (M, *)$ est désignée par le symbole Card M ou $|M|$. Lorsqu'un monoïde M comprend un nombre fini d'éléments, on dit que l'on a affaire à un monoïde fini d'ordre $|M|$.

Donnons quelques exemples de demi-groupes et de monoïdes.

1) Soient Ω un ensemble arbitraire et $M(\Omega)$ l'ensemble de toutes ses transformations (des applications de Ω dans lui-même). Les propriétés des ensembles et des applications établies au chapitre 1, § 5 entraînent que $(M(\Omega), \circ, e_\Omega)$ est un monoïde. Ici $\circ$ désigne la composition des applications et e_Ω l'application identique.

Considérons à part le cas particulier, où Ω est un ensemble fini comprenant $|\Omega| = n$ éléments désignés simplement par les entiers naturels $1, 2, \ldots, n$. Toute application $f: \Omega \to \Omega$ est définie par la donnée d'une suite ordonnée

$f(1), f(2), \ldots, f(n)$, où les images $f(i)$ sont éléments de Ω. On n'exclut pas des coïncidences $f(i) = f(j)$ pour $i \neq j$. En composant toutes les suites possibles, on obtient exactement n^n transformations. Donc, $|M(\Omega)| = \text{Card } M(\Omega) = n^n$. Posons $n = 2$ par exemple. Les éléments e, f, g, h du monoïde $M(\{1, 2\})$ et leurs produits deux à deux sont entièrement définis par deux tableaux

	1	2
e	1	2
f	2	1
g	1	1
h	2	2

$\cdot$	e	f	g	h
e	e	f	g	h
f	f	e	h	g
g	g	g	g	g
h	h	h	h	h

Il est immédiat que $M(\{1, 2\})$ est un monoïde non commutatif.

2) Soient Ω un ensemble quelconque et $\mathcal{P}(\Omega)$ l'ensemble de toutes ses parties (voir chap. 1, § 5, exercice 4). On peut munir $\mathcal{P}(\Omega)$ des opérations binaires $\cap$ et $\cup$ qui sont associatives vu que $(A \cap B) \cap C = A \cap (B \cap C)$ et $(A \cup B) \cup C = A \cup (B \cup C)$. Il est évident que $\varnothing \cup A = A$ et $A \cap \Omega = A$. On a donc deux monoïdes commutatifs $(\mathcal{P}(\Omega), \cap, \varnothing)$ et $(\mathcal{P}(\Omega), \cup, \Omega)$. Comme on le sait, $|\mathcal{P}(\Omega)| = 2^n$ si $|\Omega| = n$.

3) $(M_n(\mathbb{R}), +, 0)$ est un monoïde commutatif ayant pour élément neutre la matrice nulle, et $(M_n(\mathbb{R}), \cdot, E)$ est un monoïde non commutatif dont l'élément neutre est la matrice unité E. Cela résulte immédiatement des propriétés de l'addition et de la multiplication des matrices que nous avons étudiées au chapitre 2.

4) Soit $n\mathbb{Z} = \{nm \mid m \in \mathbb{Z}\}$ l'ensemble des entiers divisibles par n. Il est clair que $(n\mathbb{Z}, +, 0)$ est un monoïde commutatif et $(n\mathbb{Z}, \cdot)$ un demi-groupe commutatif sans élément unité $(n > 1)$.

5) L'ensemble $P_n(\mathbb{R})$ des matrices stochastiques d'ordre n (voir chap. 2, § 3, exercice 8) est un monoïde muni de l'opération de multiplication ordinaire des matrices.

Un sous-ensemble S' d'un demi-groupe S muni d'une opération $*$ s'appelle *sous-demi-groupe* si $x * y \in S'$ pour tout couple $x, y \in S'$. On dit aussi dans ce cas que le *sous-ensemble* $S' \subset S$ *est stable pour l'opération* $*$. Si $(M, *)$ est un monoïde, et le sous-ensemble $M' \subset M$ est non-seulement stable pour l'opération $*$ mais admet encore un élément neutre, on dit que M' est un *sous-monoïde* de M. Par exemple, $(n\mathbb{Z}, \cdot)$ est un sous-demi-groupe de $(\mathbb{Z}, \cdot)$, et $(n\mathbb{Z}, +, 0)$ est un sous-monoïde de $(\mathbb{Z}, +, 0)$. Tout sous-monoïde d'un monoïde $M(\Omega)$ s'appelle *monoïde de transformations* (de l'ensemble Ω).

3. Associativité généralisée; puissances.— Soit $(X, \cdot)$ une structure algébrique quelconque dont $\cdot$ est une opération binaire. Par souci de simplicité, on écrit xy au lieu de $x \cdot y$. Soit $x_1, \ldots, x_n$ une suite ordonnée d'éléments de X. Sans changer l'ordre, nous pouvons composer, par de nombreux procédés différents, des produits de longueur n. Supposons que le nombre de tels procédés soit égal

à l_n :

$l_2 = 1 :\ x_1x_2;$

$l_3 = 2 :\ (x_1x_2)\,x_3,\ x_1\,(x_2x_3);$

$l_4 = 5 :\ ((x_1x_2)\,x_3)\,x_4,\ (x_1\,(x_2x_3))\,x_4,\ x_1\,((x_2x_3)\,x_4),\ x_1\,(x_2\,(x_3x_4)),$
$(x_1,\ x_2)\,(x_3x_4)$; etc.

Il est évident qu'en prenant tous les produits possibles $x_1 \ldots x_k$, $x_{k+1} \ldots x_n$ de longueurs k et $n-k$, $1 \leqslant k \leqslant n-1$, et les composant ensuite par notre opération binaire dans l'ordre donné, nous utiliserons toutes les l_n possibilités. Il est remarquable que dans les monoïdes (et les demi-groupes) la disposition des parenthèses devient inutile.

THÉORÈME 1. — *Si une opération binaire définie sur un ensemble X est associative, le résultat de son application successive à n éléments de cet ensemble ne dépend pas de la disposition des parenthèses.*

DÉMONSTRATION. — Pour $n = 1, 2$ il n'y a rien à démontrer. Pour $n = 3$ la proposition énoncée par le théorème coïncide avec la loi d'associativité. Raisonnons plus loin par récurrence sur n. Supposons que $n > 3$ et que pour un nombre d'éléments $< n$ la vérité de la proposition soit établie. Il ne nous reste à montrer que

$$(x_1, \ldots, x_k)\cdot(x_{k+1} \ldots x_n) = (x_1 \ldots x_l)\cdot(x_{l+1} \ldots x_n) \qquad (1)$$

quels que soient k, l, $1 \leqslant k$, $l \leqslant n-1$. Nous avons mis seulement des parenthèses extérieures car par hypothèse de récurrence la disposition des parenthèses intérieures est sans importance. En particulier, $x_1x_2 \ldots x_k = (\ldots((x_1x_2)\,x_3 \ldots x_{k-1})\,x_k$ est un produit que l'on appelle produit *normé à gauche*. On distingue deux cas:

a) $k = n-1$. Alors $(x_1 \ldots x_{n-1})\,x_n = (\ldots(x_1x_2) \ldots x_{n-1})\,x_n$ est un produit normé à gauche;

b) $k < n-1$. En raison de l'associativité on a

$$(x_1 \ldots x_k)\,(x_{k+1} \ldots x_n) = (x_1 \ldots x_k)\,((x_{k+1} \ldots x_{n-1})\,x_n) =$$
$$= ((x_1 \ldots x_k)\,(x_{k+1} \ldots x_{n-1}))\,x_n =$$
$$= (\ldots((\ldots(x_1x_2) \ldots x_k)\,x_{k+1}) \ldots x_{n-1})\,x_n,$$

qui est encore un produit normé à gauche. Le second membre de l'égalité (1) à démontrer se ramène à cette même forme. ∎

Au chapitre 2, § 2 nous avons introduit le signe de sommation $\sum x_i$. Il est évident qu'on peut l'employer aussi dans tout monoïde additif commutatif. Pour un monoïde multiplicatif son analogue est le signe de produit:

$$\prod_{i=1}^{2} x_i = x_1\,x_2,\quad \prod_{i=1}^{3} x_i = (x_1x_2)\,x_3,\quad \prod_{i=1}^{n} x_i = \Big(\prod_{i=1}^{n-1} x_i\Big)\,x_n.$$

Par suite du théorème 1, dans l'écriture (et lors du calcul) du produit $x_1x_2 \ldots x_n$ des éléments d'un monoïde on peut supprimer les parenthèses. On doit avoir le seul soin de respecter l'ordre des facteurs, et cela surtout dans le cas où ils ne sont pas tous permutables. En particulier, pour $x_1 = x_2 = \ldots = x_n = x$, le produit $xx \ldots x$ est désigné, de même que lors des opérations sur les nombres, par le symbole x^n que l'on appelle *puissance n-ième de l'élément x*. Comme corollaire du théorème 1 on a les relations suivantes

$$x^m x^n = x^{m+n}, \quad (x^m)^n = x^{mn}, \quad m, n \in \mathbb{N}. \tag{2}$$

Lorsqu'il s'agit d'un monoïde $(M, \cdot, e)$ on pose encore $x^0 = e$ pour tout $x \in M$.

Etant donné les monoïdes $(M, \cdot, e)$ et $(M, +, 0)$, on associe aux puissances $x^n \in (M, \cdot, e)$ les éléments $nx = x + x + \ldots + x \in (M, +, 0)$ appelés *multiples* de x. Les règles (2) deviennent valables pour les multiples:

$$mx + nx = (m + n)\, x, \quad n\,(mx) = (nm)\, x. \tag{2'}$$

On note encore un fait utile. Si $xy = yx$ dans un monoïde M, alors

$$(xy)^n = x^n y^n, \quad n = 0, 1, 2 \ldots \tag{3}$$

En particulier, cette relation est toujours vérifiée dans un monoïde commutatif. On la démontre par récurrence sur n:

$$(xy)^n = (xy)^{n-1}(xy) = (x^{n-1}y^{n-1})(xy) = (x^{n-1}y^{n-1}x)\, y = \\ = (x^{n-1}xy^{n-1})\, y = (x^{n-1}x)(y^{n-1}y) = x^n y^n.$$

Plus généralement en s'appuyant sur la relation (3) et en utilisant la récurrence sur m, on obtient

$$x_i x_j = x_j x_i, \quad 1 \leqslant i, j \leqslant m \Rightarrow (x_1 \ldots x_m)^n = x_1^n \ldots x_m^n. \tag{4}$$

De façon analogue

$$n\,(x + y) = nx + ny, \quad n = 0, 1, 2, \ldots \tag{3'}$$

$$n\,(x_1 + \ldots + x_m) = nx_1 + \ldots + nx_m, \quad n = 0, 1, 2, \ldots \tag{4'}$$

Généralement, le monoïde $(M, \cdot, e)$ est dit *multiplicatif* et le monoïde $(M, +, 0)$ *additif*. La notation additive est réservée de préférence pour les monoïdes commutatifs.

4. Eléments inversibles.— Un élément a d'un monoïde $(M, \cdot, e)$ est dit *inversible* (*symétrisable*) s'il existe un élément $b \in M$ tel que $ab = e = ba$ (il va de soi que l'élément b sera, lui aussi, inversible). Si de plus $ab' = e = b'a$, on a $b' = eb' = (ba)\, b' = b\,(ab') = be = b$. Cela nous permet de parler tout simplement d'un *élément a^{-1} inverse* (*symétrique*) de l'élément (inversible) $a \in M$: $a^{-1}a = e = aa^{-1}$.

Bien entendu, $(a^{-1})^{-1} = a$. La notion d'élément inversible d'un monoïde constitue manifestement une généralisation naturelle de la notion de matrice inversible dans un monoïde multiplicatif $(M_n(\mathbb{R}), \cdot, E)$.

Puisque $(xy)(y^{-1}x^{-1}) = x(yy^{-1})x^{-1} = xex^{-1} = xx^{-1} = e$ et, de façon analogue, $(y^{-1}x^{-1})(xy) = e$, on a $(xy)^{-1} = y^{-1}x^{-1}$. Par conséquent, *l'ensemble de tous les éléments inversibles d'un monoïde* $(M, \cdot, e)$ *est stable pour l'opération* $\cdot$ *et forme un sous-monoïde de* M.

EXERCICES

1. Au numéro 2, nous avons introduit sur $\mathbb{Z}$, à titre d'exemple, une opération $*$: $n*m = -n - m$ qui est commutative mais non associative. Les éléments du système algébrique $(\mathbb{Z}, *)$ vérifient les relations: $(n*m)*m = n$, $m*(m*n) = n$. Soit donné maintenant un système algébrique quelconque $(X, *)$ dans lequel $(x*y)*y = x$, $y*(y*x) = x$ quels que soient $x, y \in X$. Démontrer que $x * y = y * x$, c'est-à-dire que l'opération $*$ est commutative. (Aucune indication pour la résolution, parce que c'est un exercice des plus inutiles dans le livre. Mais quand même!)

2. Montrer que

$$M_n^\circ(\mathbb{R}) = \left\{A = (a_{ij}) \in M_n(\mathbb{R}) \;\middle|\; \sum_{j=1}^{n} a_{ij} = 0,\ i = 1, 2, \ldots, n\right\}$$

muni de l'opération ordinaire de multiplication des matrices est un demi-groupe. $(M_n^\circ(\mathbb{R}), \cdot)$ est-il un monoïde?

3. Dans un monoïde multiplicatif M on choisit un élément t quelconque et on définit une nouvelle opération $*$: $x * y = xty$. Montrer que $(M, *)$ est un demi-groupe et que l'inversibilité de l'élément t de M est une condition nécessaire et suffisante pour que $(M, *)$ soit un monoïde possédant un élément neutre (élément unité) t^{-1}.

4. Montrer que l'ensemble $\mathbb{Z}$ muni de l'opération $\circ$: $n \circ m = n + m + nm = (1 + n)(1 + m) - 1$ est un monoïde commutatif. Quel est l'élément neutre de $(\mathbb{Z}, \circ)$? Indiquer dans $(\mathbb{Z}, \circ)$ tous les éléments symétrisables.

§ 2. Groupes

1. Définition et exemples.— Considérons l'ensemble $GL(n, \mathbb{R})$ de toutes les matrices carrées d'ordre n à coefficients réels et à déterminant non nul. En vertu du théorème 5, chap. 3, § 2, $\det A \neq 0$, $\det B \neq 0 \Rightarrow \det AB \neq 0$, donc $A, B \in GL(n, \mathbb{R}) \Rightarrow AB \in GL(n, \mathbb{R})$. On a $(AB)C = A(BC)$ et il existe une matrice particulière E telle que $AE = EA = A$ pour toute $A \in GL(n, \mathbb{R})$. En outre, chaque matrice $A \in GL(n, \mathbb{R})$ possède une matrice inverse A^{-1} telle que $AA^{-1} = A^{-1}A = E$.

L'ensemble $GL(n, \mathbb{R})$ muni d'une loi de composition (d'une opération binaire) $(A, B) \mapsto AB$ et appelé *groupe linéaire complet d'ordre n sur* $\mathbb{R}$, pourrait être défini tout court en utilisant la terminologie du § 1, comme *sous-monoïde de tous les éléments inversibles* du monoïde $(M_n(\mathbb{R}), \cdot, E)$. Etant d'une importance particulière, il mérite une appellation spéciale et offre un bon prétexte pour introduire une définition générale.

DÉFINITION. — *Un monoïde G dont tous les éléments sont inversibles s'appelle un groupe.*

En d'autres termes, l'ensemble G s'appelle *groupe* s'il satisfait aux axiomes suivants:

(G1) *l'ensemble G est muni d'une opération binaire (d'une loi de composition)*: $(x, y) \mapsto xy$;

(G2) *cette opération est associative*: $(xy)z = x(yz)$ *pour tout triplet* $x, y, z \in G$;

(G3) *G possède un élément neutre (élément unité)* e: $xe = ex = x$ *pour tout* $x \in G$;

(G4) *pour tout élément $x \in G$ il existe un élément inverse (élément symétrique)* x^{-1}: $xx^{-1} = x^{-1}x = e$.

Il paraît étonnant que les quatre axiomes si simples servent de base d'une branche d'algèbre, qui est parmi les plus vieilles et les plus riches en résultats, et qui joue un rôle fondamental en géométrie et dans les applications des mathématiques aux sciences naturelles. Une petite analyse montre que ces axiomes peuvent être encore simplifiés, mais ce problème ne présente pas pour nous de l'importance.

Un groupe dont la loi de composition est commutative s'appelle naturellement groupe commutatif. Un groupe commutatif est encore souvent appelé groupe *abélien* (en l'honneur du mathématicien norvégien Abel). Le terme « groupe », lui-même, a été proposé par le mathématicien français Galois qui est le vrai créateur de la théorie des groupes. Les idées de la théorie des groupes « étaient dans l'air » (comme cela arrive souvent aux idées mathématiques fondamentales) bien avant Galois, et certains de ses théorèmes ont été démontrés sous une forme naïve encore par Lagrange. Les contemporains de Galois n'ont pas compris et apprécié ses travaux géniaux. On ne s'y intéresse qu'après la parution en 1870 du livre de Jordan « Traité des substitutions et des équations algébriques ». C'est seulement vers la fin du XIXe siècle que dans la théorie des groupes « la fantaisie a été définitivement abandonnée pour faire place à une préparation soigneuse du squelette logique » (F. Klein « Conférences sur le développement des mathématiques au XIXe siècle »).

Pour désigner le nombre d'éléments dans un groupe G (plus exactement, la puissance du groupe) on utilise les symboles équivalents Card G, $|G|$ et $(G:e)$. Presque tout ce que nous venons de dire au § 1 sur les monoïdes s'applique aux groupes. Il reste seulement de remplacer convenablement les mots. En particulier on dit qu'un sous-ensemble $H \subset G$ est un *sous-groupe* de G, si $e \in H$; $h_1, h_2 \in H \Rightarrow h_1h_2 \in H$ et $h \in H \Rightarrow h^{-1} \in H$. Le sous-groupe $H \subset G$ est *propre* si $H \neq e$ et $H \neq G$.

Donnons quelques exemples de groupes.

1) Dans le groupe linéaire complet GL $(n, \mathbb{R})$ que nous connaissons déjà, examinons un sous-ensemble SL $(n, \mathbb{R})$ des matrices de déterminant 1 :

$$\mathrm{SL}(n, \mathbb{R}) = \{A \in \mathrm{GL}(n, \mathbb{R}) \mid \det A = 1\}.$$

Il est évident que $E \in \mathrm{SL}(n, \mathbb{R})$. Conformément aux résultats généraux obtenus au chapitre 3 on a pour les déterminants: $\det A = 1$, $\det B = 1 \Rightarrow \det AB = 1$ et $\det A^{-1} = (\det A)^{-1} = 1$. Donc, SL $(n, \mathbb{R})$ est un sous-groupe de GL $(n, \mathbb{R})$ que l'on appelle *groupe linéaire spécial d'ordre n sur* $\mathbb{R}$. Il s'appelle encore groupe *unimodulaire*, bien que souvent on appelle ainsi le groupe des matrices dont le déterminant est égal à ± 1.

Il faut dire que le groupe GL $(n, \mathbb{R})$ qui est un vrai récipient de nombreux groupes bien intéressants, constitue pour les mathématiciens de différentes générations une sorte de source intarissable de nouvelles idées et de problèmes à résoudre.

2) En utilisant les nombres rationnels au lieu des nombres réels, nous obtenons le groupe linéaire complet GL $(n, \mathbb{Q})$ d'ordre n sur $\mathbb{Q}$ et son sous-groupe SL $(n, \mathbb{Q})$. A son tour, SL $(n, \mathbb{Q})$ contient un sous-groupe intéressant

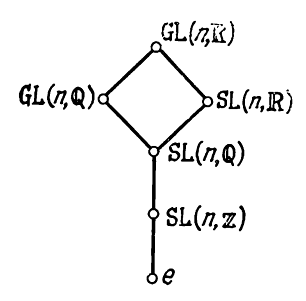


Fig. 11.

SL $(n, \mathbb{Z})$ des matrices à coefficients entiers de déterminant 1. Le théorème 1 du chapitre 3, § 3, qui donne la formule explicite pour les coefficients d'une matrice inverse, montre que SL $(n, \mathbb{Z})$ est réellement un groupe. Les groupes SL $(n, \mathbb{Q})$ et SL $(n, \mathbb{Z})$ occupent une place d'honneur dans la théorie des nombres. L'ensemble partiellement ordonné (voir chap. 1, § 6, n° 3) des sous-groupes du groupe GL $(n, \mathbb{R})$ que nous venons de considérer est représenté par le diagramme de la fig. 11.

3) En posant $n = 1$ dans les exemples 1) et 2), nous obtenons les groupes multiplicatifs $\mathbb{R}^* = \mathbb{R} \setminus \{0\} = \mathrm{GL}(1, \mathbb{R})$, $\mathbb{Q}^* = \mathbb{Q} \setminus \{0\} = \mathrm{GL}(1, \mathbb{Q})$ des nombres réels et des nombres rationnels. Ces groupes sont évidemment infinis. Les seuls éléments inversibles de $(\mathbb{Z}, \cdot, 1)$ étant 1 et -1, on a $\mathrm{GL}(1, \mathbb{Z}) = \{\pm 1\}$. Par conséquent, $\mathrm{SL}(1, \mathbb{R}) = \mathrm{SL}(1, \mathbb{Q}) = \mathrm{SL}(1, \mathbb{Z}) = 1$. Lorsque $n = 2$, le groupe SL $(2, \mathbb{Z})$ est déjà infini: il contient par exemple toutes les matrices

$$\begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ m & 1 \end{pmatrix}, \quad \begin{pmatrix} m & m-1 \\ 1 & 1 \end{pmatrix}, \quad m \in \mathbb{Z}.$$

Signalons d'autres groupes additifs infinis:

$$(\mathbb{R}, +, 0), \quad (\mathbb{Q}, +, 0), \quad (\mathbb{Z}, +, 0).$$

4) Soit Ω un ensemble quelconque et soit $S(\Omega)$ l'ensemble de toutes les bijections (transformations biunivoques) $f\colon \Omega \to \Omega$. En nous reportant aux résultats obtenus au chapitre 1, § 5 sur les applications des ensembles (théorèmes 1, 2 et corollaire du théorème 2), nous arrivons immédiatement à cette conclusion que $S(\Omega)$ est un groupe relativement à la loi de composition des applications qui est une opération binaire naturelle. Il est évident que $S(\Omega)$ est un sous-monoïde constitué de tous les éléments inversibles du monoïde $M(\Omega)$ de l'exemple 1) du § 1, mais il n'est pas dans notre intention de souligner ce fait. Le groupe $S(\Omega)$ lui-même et surtout ses divers sous-groupes appelés *groupes de transformations* constituent, si l'on peut dire ainsi, une plate-forme de départ pour toutes les applications pratiques possibles de la théorie des groupes. Il suffit de mentionner « Le programme d'Erlangen » devenu célèbre de Felix Klein (1872), qui a adopté la notion de groupe de transformations comme base pour la classification de divers types de géométries. En prenant pour Ω l'espace vectoriel $\mathbb{R}^n$, on obtient un « grand » groupe $S(\mathbb{R}^n)$ qui n'est pas facile à passer en revue et qui contient un sous-groupe des applications linéaires inversibles (bijectives) $\varphi_A : \mathbb{R}^n \to \mathbb{R}^n$ qui sont en correspondance biunivoque avec les matrices régulières A d'ordre n (voir chap. 2, § 3).

Ainsi, on obtient une injection de $\mathrm{GL}(n, \mathbb{R})$ dans $S(\mathbb{R}^n)$. Le sens de cette injection deviendra plus clair quand nous introduirons la notion importante d'isomorphisme de groupes.

2. Système de générateurs (lors d'une première lecture ce numéro peut être omis).

Etant donné un sous-ensemble S du groupe G, proposons-nous de trouver un sous-groupe $H \subset G$ contenant S et tel que pour tout sous-groupe $K \subset G$, $S \subset K$ entraîne $H \subset K$. Il ne peut pas exister deux sous-groupes minimaux H, H' de ce genre; en effet

$$S \subset H,\ S \subset H' \Rightarrow H \subset H' \subset H \Rightarrow H' = H.$$

Ainsi, le sous-groupe minimal H doit coïncider avec l'intersection de tous les sous-groupes contenant S, si seulement cette intersection est un sous-groupe de G. On a un théorème simple:

Théorème 1. — *L'intersection* $\bigcap_{i \in I} H_i$ *de toute famille* $\{H_i \mid i \in I\}$ *de sous-groupes d'un groupe G est un sous-groupe.*

Démonstration. — Soit e l'élément unité du groupe G. Les propriétés $e \in \bigcap H_i$; $x, y \in \bigcap H_i \Rightarrow xy \in \bigcap H_i$; $x \in \bigcap H_i \Rightarrow x^{-1} \in \in \bigcap H_i$, qui caractérisent tout sous-groupe, sont vérifiées dans $\bigcap H_i$ parce qu'elles le sont dans chacun des sous-groupes H_i. ■

Prenons maintenant pour famille $\{H_i \mid i \in I\}$ tous les sous-groupes qui contiennent le sous-ensemble donné $S \subset G$. Leur intersection

$$\langle S \rangle = \bigcap_{S \subseteq H} H$$

sera, par suite du théorème 1 et des remarques faites ci-dessus, justement le sous-groupe minimal contenant S. Nous dirons que: $\langle S \rangle$ est un sous-groupe *engendré* par l'ensemble S dans le groupe G et que S est l'ensemble des *générateurs* du sous-groupe $\langle S \rangle$. Il semble à première vue que $\langle S \rangle$ est défini de façon inefficace, car il faut

examiner tous les sous-groupes contenant S. Pourtant, il n'est pas nécessaire de le faire, comme le montre une assertion simple qui découle du théorème 1:

COROLLAIRE.— *Le sous-groupe* $\langle S \rangle \subset G$ *coïncide avec l'ensemble* T *contenant un élément unité* e *et tous les produits possibles*

$$t_1 t_2 \ldots t_n, \quad n = 1, 2, 3, \ldots,$$

où soit $t_i \in S$, *soit* $t_i^{-1} \in S$, $1 \leqslant i \leqslant n$.

En effet, puisque $t_1 \ldots t_n \in T$, $t'_1 \ldots t'_m \in T \Rightarrow t''_1 \ldots t''_{n+m} = t_1 \ldots t_n t'_1 \ldots t'_m \in T$ et $t_1 \ldots t_n \in T \Rightarrow (t_1 \ldots t_n)^{-1} = t_n^{-1} \ldots t_1^{-1} \in T$, l'ensemble T est sous-groupe de G. D'autre part, chaque sous-groupe H, contenant tous les $x_i \in S$, doit aussi contenir tous les éléments inverses x_i^{-1} et donc tous les produits de la forme $t_1 t_2 \ldots t_n$. Par suite, $H \supset T$ et T coïncide avec l'intersection de tous ces sous-groupes. ■

Il est à remarquer que tous les produits $t_1 t_2 \ldots t_n$ sont loin d'être des éléments distincts du sous-groupe $\langle S \rangle$ même si l'on convient (ce qui est naturel) de remplacer les produits aa^{-1}, $a^{-1}a$ par l'élément unité. Dans le cas où $|S| > 1$, la question relative à l'égalité des produits $t_1 t_2 \ldots t_n$ est assez délicate à résoudre et ne sera étudiée sommairement qu'au chapitre 7.

Chaque groupe G est engendré par un système de générateurs S quelconque: l'ensemble S peut par exemple coïncider avec tout le groupe G. Pour simplifier l'étude, considérons un groupe G *engendré par un ensemble fini* S de ses éléments. En rejetant de S les éléments « superflus » qui s'écrivent sous la forme d'un produit des éléments restants (et de leurs inverses), nous obtenons un système *minimal* de générateurs M du groupe G. Cela signifie que $\langle M \rangle = G$ et que $\langle M' \rangle \neq G$ si le système M' est obtenu à partir de M en supprimant au moins un élément. Soit $M = \{g_1, \ldots, g_d\}$. Alors, au lieu de $G = \langle M \rangle$, on peut écrire $G = \langle g_1, g_2, \ldots, g_d \rangle$. Si $d = 1$, on dit que le groupe est *cyclique*.

3. Groupes cycliques.— Si G est un groupe quelconque et g son élément, alors $\langle g \rangle$ est, par définition, un sous-groupe cyclique de G.

Conformément au théorème 1 et aux propriétés des puissances des éléments dans les monoïdes, il est naturel de s'attendre que tout groupe cyclique $\langle a \rangle$ de générateur a soit un groupe abélien de la forme $\langle a \rangle = \{a^n \mid n \in \mathbb{Z}\}$ ou $\langle a \rangle = \{na \mid n \in \mathbb{Z}\}$ suivant que le groupe considéré est multiplicatif ou additif (cette notation ne signifie pas que tous les éléments a^n ou na sont distincts). Il en est justement ainsi. Pour le montrer il suffit d'adopter la notation $(a^{-1})^k = a^{-k}$ et prouver l'assertion suivante:

THÉORÈME 2.— *Quels que soient* m, $n \in \mathbb{Z}$, *on a*

$$a^m a^n = a^{m+n}, \quad (a^m)^n = a^{mn}$$

(*respectivement* $ma + na = (m + n)\, a$, $n\,(ma) = (nm)\, a$).

DÉMONSTRATION. — Pour m, n non négatifs, voir les relations (2), (2′) du § 1, n° 4. Si $m < 0$, $n < 0$, on a $m' = -m > 0$, $n' = -n > 0$ et

$$a^m a^n = (a^{-1})^{m'} (a^{-1})^{n'} = (a^{-1})^{m'+n'} = a^{-(m'+n')} = a^{m+n}.$$

Pour $m' = -m > 0$, $n > 0$ on a

$$a^m a^n = (a^{-1})^{m'} a^n = \underbrace{(a^{-1} \ldots a^{-1})}_{m'} \underbrace{(a \ldots a)}_{n} =$$

$$= a^{n-m'} \text{ (ou } (a^{-1})^{m'-n}, \text{ si } m' \geqslant n) = a^{m+n}.$$

On considère de façon analogue le cas où $m > 0$, $n < 0$. L'égalité $(a^m)^n = a^{mn}$ découle de ce qui précède et devient suffisamment évidente si l'on se reporte à la définition des puissances. ■

L'exemple le plus simple de groupe cyclique est fourni par le groupe additif des entiers $(\mathbb{Z}, +, 0)$ engendré par 1 ou -1. On vérifie aussi aisément que la matrice $\begin{Vmatrix} 1 & 1 \\ 0 & 1 \end{Vmatrix}$ engendre dans $SL(2, \mathbb{Z})$ un sous-groupe cyclique infini. L'ensemble $\{1, -1\}$ forme un groupe cyclique d'ordre 2 pour l'opération de multiplication.

On obtient un exemple de groupe cyclique d'ordre n en considérant toutes les rotations sur le plan autour d'un point O, qui amènent en coïncidence avec lui-même un polygone régulier P_n de n côtés centré en O. Il est évident que ces rotations forment un groupe relativement à la loi de composition qui est une réalisation successive de transformations. Notre groupe C_n contient les rotations $\varphi_0, \varphi_1, \ldots, \varphi_{n-1}$ d'angles $0, \frac{2\pi}{n}, \ldots, (n-1)\frac{2\pi}{n}$ effectuées en sens inverse des aiguilles d'une montre. Ceci étant $\varphi_s = \varphi_1^s$. Une interprétation géométrique des rotations permet de dire que $\varphi_s^{-1} = \varphi_1^{n-s}$ et $\varphi_1^n = \varphi_0$ (transformation identique). Ainsi, $|C_n| = n$ et $C_n = \langle \varphi_1 \rangle$. Remarquons que le groupe cyclique C_n est un sous-groupe propre du groupe D_n de toutes les transformations de symétrie du polygone P_n de n côtés (c'est-à-dire des coïncidences de P_n avec lui-même).

Soient G un groupe quelconque et a un élément de G. Il y a deux possibilités: 1) Toutes les puissances de l'élément a sont distinctes, c'est-à-dire $m \neq n \Rightarrow a^m \neq a^n$. On dit dans ce cas que l'élément $a \in G$ est d'*ordre infini*. 2) Il a y des coïncidences $a^m = a^n$ pour $m \neq n$. Si, par exemple, $m > n$, on a $a^{m-n} = e$, c'est-à-dire il existe des puissances positives de l'élément $a \in G$ égales à l'élément unité. Soit q le plus petit exposant positif pour lequel $a^q = e$. On dit alors que a est un élément d'*ordre fini* q. Dans un groupe fini G (Card $G < \infty$), tous les éléments sont évidemment d'ordre fini.

AVERTISSEMENT.— En mathématiques, le mot « ordre » a plusieurs sens. Nous avons rencontré précédemment les matrices carrées d'ordre n (les matrices $n \times n$). Or une matrice régulière A considérée comme élément du groupe GL $(n, \mathbb{R})$ a aussi un ordre (peut-être infini) dans le sens que nous venons d'indiquer. Chaque fois le contexte permet de comprendre de quoi il s'agit.

Sur le fond de l'exemple sus-indiqué de groupe cyclique d'ordre n, l'assertion suivante est presque évidente.

THÉORÈME 3. — *L'ordre de tout élément $a \in G$ (G est un groupe quelconque) est égal à* Card $\langle a \rangle$. *Si a est un élément d'ordre fini q, on a*

$$\langle a \rangle = \{e, a, \ldots, a^{q-1}\} \quad \text{et} \quad a^k = e \Leftrightarrow k = lq, \quad l \in \mathbb{Z}.$$

DÉMONSTRATION. — Dans le cas d'un élément d'ordre infini il n'y a rien à démontrer. Si a est d'ordre q alors par définition tous les éléments $e, a, a^2, \ldots, a^{q-1}$ sont distincts. Toute autre puissance a^k coïncide avec l'un de ces éléments, c'est-à-dire $\langle a \rangle = \{e, a, \ldots$ $\ldots, a^{q-1}\}$. En effet, utilisons l'algorithme de division dans $\mathbb{Z}$ (chap. 1, § 8, n° 3) et écrivons l'exposant k sous la forme

$$k = lq + r, \quad 0 \leqslant r \leqslant q - 1.$$

Après quoi, en appliquant aux puissances les règles énoncées par le théorème 2, on obtient

$$a^k = (a^q)^l a^r = ea^r = a^r.$$

En particulier, $a^k = e \Rightarrow r = 0 \Rightarrow k = lq$. ■

La propriété d'un groupe d'être cyclique est très utile et commode, mais elle n'est pas toujours donnée à priori. Parfois on est amené à la démontrer. Considérons à titre d'exemple la proposition suivante :

PROPOSITION. — *Les éléments permutables a, b d'un groupe quelconque G, ayant des ordres s, t premiers entre eux, engendrent dans G un sous-groupe cyclique d'ordre st*

$$\langle a, b \rangle = \langle ab \rangle.$$

DÉMONSTRATION. — En effet, $D = \langle a \rangle \cap \langle b \rangle = e$ car en vertu du théorème 3 pour tout élément $d \in D$ d'ordre q quelconque, on a

$$d = a^i = b^j \Rightarrow d^s = (a^s)^i = e, \; d^t = (b^t)^j = e \Rightarrow q \mid s, \; q \mid t;$$

s et t étant premiers entre eux, il résulte $q = 1$. Soit $n = |\langle ab \rangle|$. On a (voir § 1, relation (3)) :

$$a^n b^n = (ab)^n = e \Rightarrow a^n = b^{-n} \in D = e \Rightarrow a^n = e, \qquad b^n = e \Rightarrow$$
$$\Rightarrow s \mid n, \; t \mid n \Rightarrow \text{P.P.C.M.}\ (s, t) \mid n \Rightarrow st \mid n,$$

puisque $st =$ P.P.C.M. (s, t) P.G.C.D. (s, t). Or, $(ab)^{st} = (a^s)^t (b^t)^s =$ $= e$ (théorème 2), si bien que $n \mid st$ et donc $n = st$. Il reste à remar-

quer que

$$\langle a, b\rangle = \{a^i b^j \mid 0 \leqslant i \leqslant s-1, \quad 0 \leqslant j \leqslant t-1\} \Longrightarrow \mathrm{Card}\,\langle a, b\rangle \leqslant st.$$

Vu que $\langle ab\rangle \subset \langle a, b\rangle$ et Card $\langle ab\rangle = st$, on a $\langle a, b\rangle = \langle ab\rangle$. ■

Nous reviendrons plus tard sur les groupes cycliques. Procédons maintenant à une étude approfondie d'un type spécial de groupes de transformations qui permettent de mieux mettre en évidence les notions que nous avons introduites.

4. Groupe symétrique et groupe alterné.— Soit Ω un ensemble fini à n éléments. La nature de ses éléments étant pour nous sans importance, il est commode de considérer $\Omega = \{1, 2, \ldots, n\}$. Le groupe $S(\Omega)$ (voir plus haut l'exemple 4) de toutes les applications bijectives $\Omega \to \Omega$ s'appelle *groupe symétrique de degré n* (ou encore *groupe symétrique à n symboles* ou *à n points*), et se note le plus souvent S_n.

Ses éléments, désignés généralement par des lettres minuscules de l'alphabet grec s'appellent *permutations.*

REMARQUE. — Autrefois (et parfois actuellement) les éléments du groupe S_n étaient appelés « substitutions », le mot « permutation » désignant alors un arrangement des nombres 1, 2, . . ., n dans un ordre fixe quelconque. Puisque les arrangements de nombres sont en correspondance biunivoques avec les éléments du groupe S_n, et le mot « permutation » s'associe dans l'esprit plutôt à l'action qu'à un arrangement figé, le terme « substitution » a été abandonné. D'ailleurs, nous parlerons plus loin des substitutions des nombres dans un polynôme, mais ce n'est qu'un argument de plus en faveur de la convention terminologique adoptée.

Sous une forme développée et bien suggestive la permutation $\pi: i \mapsto \pi(i)$, $i = 1, 2, \ldots, n$ est représentée par le symbole

$$\pi = \begin{pmatrix} 1 & 2 \ldots n \\ i_1 & i_2 \ldots i_n \end{pmatrix},$$

où l'on indique *in extenso* toutes les images

$$\pi:\ \begin{matrix} 1 & 2 & \ldots & n \\ \downarrow & \downarrow & & \downarrow \\ i_1 & i_2 & \ldots & i_n \end{matrix},$$

$i_k = \pi(k)$, $k = 1, \ldots, n$ étant les symboles permutés 1, 2, ..., n. Comme toujours e est une permutation identique (bien que e soit une lettre d'alphabet latin) : $e(i) = i$, $\forall i$.

Les permutations $\sigma, \tau \in S_n$ sont multipliées suivant la loi de composition des applications : $(\sigma\tau)(i) = \sigma(\tau(i))$. C'est ainsi par exemple que pour les permutations

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}, \quad \tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix}$$

on a

$$\sigma\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix} = \begin{matrix} 1 & 2 & 3 & 4 \\ \downarrow & \downarrow & \downarrow & \downarrow \\ 4 & 3 & 2 & 1 \\ \downarrow & \downarrow & \downarrow & \downarrow \\ 1 & 4 & 3 & 2 \end{matrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix}.$$

Etant donné

$$\tau\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix},$$

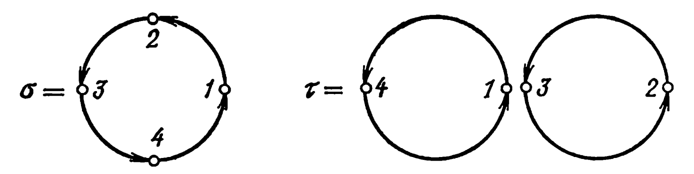
il vient $\sigma\tau \neq \tau\sigma$.

Parfois, à côté du groupe G, il est commode de considérer un groupe dit *opposé*. Si G est un groupe pour une opération binaire $\circ: (f, g) \mapsto f \circ g$, il est aussi un groupe pour une opération $*$: $(f, g) \mapsto g \circ f$. Le groupe muni de l'opération opposée se note G^{op}. Ce fait reflète la symétrie des axiomes de groupes, où il s'agit des éléments symétriques bilatères et des éléments neutres bilatères. L'axiome d'associativité est aussi symétrique. En particulier, dans le groupe S_n^{op} on définit la règle de multiplication de deux permutations dans le sens ordinaire de gauche à droite. Si nous avions écrit $i\sigma$ ou i^σ au lieu de $\sigma i = \sigma(i)$ ce serait bien habituel aussi pour nous.

Cherchons maintenant à définir l'ordre du groupe S_n. Par une permutation σ le symbole 1 peut être transformé en un $\sigma(1)$ et il existe à cet effet exactement n possibilités différentes. Après avoir fixé $\sigma(1)$ nous avons le droit de prendre pour $\sigma(2)$ seulement l'un des $n - 1$ symboles restants (il existe au total $(n - 1) + (n - 1) + \ldots + (n - 1) = n(n - 1)$ couples dictincts $\sigma(1), \sigma(2)$), pour $\sigma(3)$ respectivement $n - 2$ symboles, etc. On a au total $\sigma(1)$, $\sigma(2), \ldots, \sigma(n)$ possibilités de choix et donc le nombre de toutes les permutations différentes est égal à $n(n - 1) \ldots 2\cdot 1 = n!$ (factorielle n). Ainsi

$$\text{Card } S_n = |S_n| = (S_n : e) = n!.$$

Décomposons maintenant les permutations de S_n en un produit de permutations plus simples. L'idée de cette décomposition sera expliquée schématiquement sur l'exemple des permutations σ, $\tau \in S_4$ indiquées plus haut :



La permutation σ écrite en abrégé sous la forme $\sigma = (1234)$ ou, ce qui revient au même, sous la forme $\sigma = (2341) = (3412) = (4123)$ s'appelle *cycle* de longueur 4, alors que la permutation $\tau = (14)\ (23)$

s'appelle produit de deux cycles *indépendants* (*de supports disjoints*) (14) et (23) de longueur 2. Remarquons que $\sigma^2 = (13)\,(24)$, $\sigma^4 = (\sigma^2)^2 = e$, $\tau^2 = e$.

En passant au cas général, nous dirons que deux points $i, j \in \Omega$ sont *équivalents* par rapport au sous-groupe cyclique $\langle\pi\rangle \subset S_n$ ou tout simplement π-*équivalents* si $j = \pi^s(i) = \pi(\ldots\pi(i)\ldots)$ pour un $s \in \mathbb{Z}$. Puisque S_n est un groupe fini, chacun de ses sous-groupes est aussi fini. Etant donné que Card $\langle\pi\rangle = q$, on peut poser $0 \leqslant s < q$ (théorème 3). Nous avons affaire à une relation réflexive, symétrique et transitive (voir chap. 1, § 6, n° 2) car $i = \pi^0(i) = e(i)$; $j = \pi^k(i) \Rightarrow i = \pi^{q-k}(j)$ et $j = \pi^s(i)$, $k = \pi^t(j) \Rightarrow k = \pi^{s+t}(i)$. Conformément à la propriété générale des relations d'équivalence, on obtient la partition

$$\Omega = \Omega_1 \cup \ldots \cup \Omega_p \qquad (1)$$

de l'ensemble Ω en classes disjointes deux à deux $\Omega_1, \ldots, \Omega_p$ qu'on convient d'appeler π-*orbites*. Cette appellation est parfaitement justifiée. Chaque point $i \in \Omega$ appartient exactement à une orbite, et si $i \in \Omega_k$, alors Ω_k se forme des images de i par les applications des puissances de l'élément π, à savoir: $i, \pi(i), \pi^2(i), \ldots, \pi^{l_k-1}(i)$. Ici, $l_k = |\Omega_k|$ est la longueur de la π-orbite Ω_k. Il est évident que $l_k \leqslant q =$ Card $\langle\pi\rangle$ et $\pi^{l_k}(i) = i$, l_k étant le plus petit nombre possédant cette propriété. Posant

$$\pi_k = (i\pi(i)\ldots\pi^{l_k-1}(i)) = \begin{pmatrix} i & \pi(i) & \ldots & \pi^{l_k-2}(i) \\ \pi(i) & \pi^2(i) & \ldots & \pi^{l_k-1}(i) \end{pmatrix},$$

nous retrouvons justement la permutation appelée cycle de longueur l_k.

On écrit aussi bien $(123\ldots l)$ que $(1, 2, 3, \ldots, l)$ en séparant les symboles par les virgules. Le cycle π_k laisse invariants tous les points de l'ensemble $\Omega \setminus \Omega_k$, et $\pi(j) = \pi_k(j)$ pour tout point $j \in \Omega_k$. Cette propriété nous permet d'appeler π_s, π_t, $s \neq t$ cycles *indépendants* ou cycles *disjoints*. Puisque $\pi^{l_k}(i) = i$ pour tout $i \in \Omega_k$, on a $\pi_k^{l_k} = e$.

Ainsi, à la partition (1) est associée la décomposition de la permutation π en un produit

$$\pi = \pi_1\pi_2\ldots\pi_p, \qquad (2)$$

où tous les cycles sont permutables: $\pi = \pi_1\pi_2\ldots\pi_p = \pi_{i_1}\pi_{i_2}\ldots\pi_{i_p}$.

On peut poser par exemple que $l_1 \geqslant l_2 \geqslant \ldots \geqslant l_m > l_{m+1} = \ldots = l_p = 1$.

Si le cycle $\pi_k = (i)$ est de longueur 1, il opère comme une permutation identique; il est naturel d'omettre de tels cycles dans le pro-

duit (2):

$$\pi = \pi_1\pi_2 \ldots \pi_m; \quad l_k > 1, \quad 1 \leqslant k \leqslant m. \tag{3}$$

Par exemple, la permutation

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 3 & 4 & 5 & 1 & 7 & 6 & 8 \end{pmatrix} \in S_8$$

sera écrite sous la forme

$$\pi = (12345)\ (67)\ (8) = (12345)\ (67). \tag{4}$$

Un certain inconvénient peut provenir du fait que (12345) (67) peut être interprété comme une permutation de S_n pour tout $n \geqslant 7$. Or, lorsque n est fixé, toute indétermination disparaît.

D'une manière plus précise, soit donnée en plus de la décomposition (3) encore une décomposition $\pi = \alpha_1\alpha_2 \ldots \alpha_r$ en un produit de cycles indépendants, et soit i un point qui ne reste pas fixe lors de la permutation π. Alors, $\pi_s(i) \neq i$, $\alpha_t(i) \neq i$ pour l'un (et un seul) des cycles $\pi_1, \ldots, \pi_m$ et l'un des $\alpha_1, \ldots, \alpha_r$. On a

$$\pi_s^k(i) = \pi^k(i) = \alpha_t^k(i), \quad k = 0, 1, 2, \ldots$$

Or, un cycle est défini univoquement par les valeurs que ses puissances prennent en un point qui ne reste pas fixe. Par suite, $\pi_s = \alpha_t$. On continue la démonstration en raisonnant par récurrence sur m ou sur r.

Ainsi nous avons démontré le théorème suivant:

THÉORÈME 4. — *Toute permutation $\pi \neq e$ de S_n se décompose en un produit de cycles indépendants de longueur $\geqslant 2$, la décomposition étant définie univoquement à un ordre de succession des cycles près.* ■

La notation compacte (3) de la permutation π dont il s'agit dans le théorème 4, s'avère commode pour plusieurs raisons. En particulier, elle permet de déterminer sans difficulté l'ordre d'une permutation.

COROLLAIRE 1. — *L'ordre de la permutation $\pi \in S_n$ (= ordre du sous-groupe cyclique $\langle\pi\rangle$) est égal au plus petit commun multiple des longueurs des cycles indépendants intervenant dans la décomposition de* π (voir la proposition à la fin du nº 3).

DÉMONSTRATION. — Nous avons déjà indiqué que dans la décomposition $\pi = \pi_1\pi_2 \ldots \pi_m$ les cycles indépendants sont permutables, ou comme on le dit encore, commutent. C'est pourquoi, d'après la relation (4) du § 1 on a

$$\pi^s = \pi_1^s \ldots \pi_m^s, \quad s = 0, 1, 2, \ldots$$

Les cycles $\pi_1, \ldots, \pi_m$ étant indépendants (ils opèrent sur des ensembles différents $\Omega_1, \ldots, \Omega_m$) on a $\pi^q = e \Leftrightarrow \pi_q^k = e$ pour $k = 1, \ldots, m$. Donc, q est un multiple commun des ordres des cycles π_k qui coïncident, comme nous l'avons vu, avec leurs longueurs l_k. Si q est le plus petit entier naturel pour lequel $\pi^q = e$, alors $q = \text{Card}\,\langle\pi\rangle$ et $q = \text{P.P.C.M.}\ (l_1, \ldots, l_m)$ est l'entier défini au chapitre 1, § 8, n° 2. ■

Par exemple, nous pouvons dire tout de suite que l'ordre de la permutation de la forme (4) est égal à 10. Mais quel est l'ordre maximal des éléments de S_8? Examinons toutes les présentations du nombre 8 sous forme de somme des termes positifs disposés dans l'ordre non décroissant. Il vient que les ordres des éléments $\neq e$ de S_8 sont les entiers 2, 3, 4, 5, 6, 7, 8, 10, 12, 15. Pour élément d'ordre maximal 15 on peut prendre par exemple la permutation $\pi = (12345)(678)$.

Définition. — *On appelle transposition un cycle de longueur* 2.

Toute transposition est de la forme $\tau = (ij)$ et laisse fixes tous les symboles différents de i, j. Le théorème 4 entraîne le corollaire suivant:

Corollaire 2. — *Toute permutation $\pi \in S_n$ est un produit de transpositions.*

En effet, en raison du théorème 4, il suffit d'écrire chacun des cycles sous la forme d'un produit de transpositions. On peut le faire par exemple ainsi:

$$(1\ 2 \ldots l-1\ l) = (1l)\,(1\,l-1) \ldots (13)\,(12). \ ■$$

L'assertion du corollaire 2 peut être exprimée d'une autre façon, en utilisant la notion de système de générateurs d'un groupe (voir n° 2):

$$S_n = \langle (12), \ldots, (1n), (23), \ldots, (2n), \ldots, (n-1\,n) \rangle.$$

Bien entendu, ce système de générateurs n'est pas minimal. Par exemple

$$S_3 = \langle (12), (13), (23) \rangle = \langle (12), (13) \rangle.$$

On ne peut certes parler d'aucune unicité de présentation d'une permutation en fonction de transpositions: en général, les transpositions ne commutent pas et leur nombre n'est pas un invariant de la permutation. On a, par exemple, dans S_4

$$(123) = (13)\,(12) = (23)\,(13) = (13)\,(24)\,(12)\,(14).$$

D'ailleurs, la non-unicité de la décomposition découle de l'égalité $\sigma\tau^2 = \sigma$ quelles que soient les transpositions σ et τ. Il existe néanmoins un invariant de décomposition d'une permutation en des transpositions. Pour le mettre en évidence d'une manière aussi naturelle que possible, considérons le cas, où les éléments de S_n agissent sur les fonctions.

DÉFINITION. — *Soit* $\pi \in S_n$ *et soit* $f(X_1, \ldots, X_n)$ *une fonction de n arguments quelconques. Posons*

$$(\pi \circ f)(X_1, \ldots, X_n) = f(X_{\pi^{-1}(1)}, \ldots, X_{\pi^{-1}(n)}). \tag{5}$$

On dit que la fonction $g = \pi \circ f$ *est obtenue par l'action de* π *sur* f.

Par exemple si $\pi = (123)$ et $f(X_1, X_2, X_3) = X_1 + 2X_2^2 + 3X_3^3$, alors $(\pi \circ f)(X_1, X_2, X_3) = X_3 + 2X_1^2 + 3X_2^3$.

Conformément au § 1 du chap. 3, la fonction f est dite *symétrique gauche* si $\tau \circ f = -f$ pour toute transposition $\tau \in S_n$, c'est-à-dire si

$$f(\ldots, X_j, \ldots, X_i, \ldots) = -f(\ldots, X_i, \ldots, X_j, \ldots).$$

LEMME. — *Soient* α, β *deux permutations quelconques de* S_n. *Alors*

$$(\alpha\beta) \circ f = \alpha \circ (\beta \circ f).$$

DÉMONSTRATION. — D'après la relation de définition (5) on a

$$((\alpha\beta \circ f)(X_1, \ldots, X_n) = f(X_{(\alpha\beta)^{-1}(1)}, \ldots, X_{(\alpha\beta)^{-1}(n)}) =$$
$$= f(X_{(\beta^{-1}\alpha^{-1})(1)}, \ldots, X_{(\beta^{-1}\alpha^{-1})(n)}) =$$
$$= f(X_{\beta^{-1}(\alpha^{-1}1)}, \ldots, X_{\beta^{-1}(\alpha^{-1}n)}) =$$
$$= (\beta \circ f)(X_{\alpha^{-1}(1)}, \ldots, X_{\alpha^{-1}(n)}) = (\alpha \circ (\beta \circ f))(X_1, \ldots, X_n). \blacksquare$$

THÉORÈME 5. — *Soit* π *une permutation de* S_n,

$$\pi = \tau_1\tau_2 \ldots \tau_k \tag{6}$$

une décomposition quelconque de π *en un produit de transpositions. Alors le nombre*

$$\varepsilon_\pi = (-1)^k, \tag{7}$$

appelé parité de π (ou encore, *signature* ou *signe de* π) *est entièrement défini par la permutation* π *et ne dépend pas du mode de décomposition* (6), *c'est-à-dire la parité de l'entier k pour une permutation donnée* π *est toujours la même. De plus*

$$\varepsilon_{\alpha\beta} = \varepsilon_\alpha\varepsilon_\beta \tag{8}$$

quels que soient α, $\beta \in S_n$.

DÉMONSTRATION. — Prenons une fonction symétrique gauche quelconque f de n arguments $X_1, \ldots, X_n$. D'après le lemme opérer f par π c'est appliquer successivement à f les transpositions τ_k, $\tau_{k-1}, \ldots, \tau_1$, c'est-à-dire multiplier f k fois par -1 :

$$\pi \circ f = (\tau_1 \ldots \tau_{k-1}) \circ (\tau_k \circ f) =$$
$$= -(\tau_1 \ldots \tau_{k-1}) \circ f = \ldots = (-1)^k f = \varepsilon_\pi f.$$

Puisque le premier membre de cette relation dépend de π et non de l'une quelconque de ses décompositions, l'application $\varepsilon: \pi \mapsto \varepsilon_\pi$

donnée par l'égalité (7), doit se définir entièrement par la permutation π, à condition, bien sûr, que f ne soit pas une fonction identiquement nulle. Or, nous savons qu'il existe des fonctions symétriques gauches non nulles, par exemple le déterminant de Vandermonde $\Delta_n (X_1, \ldots, X_n)$ d'ordre n.

D'après la règle énoncée dans le lemme, l'application de la permutation $\alpha\beta$ à une telle fonction f donne

$$\varepsilon_{\alpha\beta} f = (\alpha\beta) \circ f = \alpha \circ (\beta \circ f) = \alpha \circ (\varepsilon_\beta f) = \\ = \varepsilon_\beta (\alpha \circ f) = \varepsilon_\beta (\varepsilon_\alpha f) = (\varepsilon_\alpha \varepsilon_\beta) f,$$

d'où l'on tire la relation (8). ■

DÉFINITION. — *Une permutation* $\pi \in S_n$ *est dite paire si* $\varepsilon_\pi = 1$ *et impaire si* $\varepsilon_\pi = -1$.

De cette définition on déduit que *toutes les transpositions sont des permutations impaires.*

COROLLAIRE 1. — *Toutes les permutations paires de degré n forment un sous-groupe* $A_n \subset S_n$ *d'ordre* $n!/2$ (il s'appelle *groupe alterné de degré* n).

DÉMONSTRATION. — D'après (8) on a $\varepsilon_{\alpha\beta} = 1$ si $\varepsilon_\alpha = \varepsilon_\beta = 1$, et $\varepsilon_{\pi^{-1}} = \varepsilon_\pi$ car $\varepsilon_e = 1$. Puisque A_n est un sous-ensemble de S_n, tous les axiomes du groupe se trouvent vérifiés.

Ecrivons S_n sous la forme de la réunion $S_n = A_n \cup \bar{A}_n$, où $\bar{A}_n$ est l'ensemble de toutes les permutations impaires de degré n. L'application de S_n dans lui-même, définie par la loi

$$\rho_{(12)} : \pi \mapsto (12)\, \pi,$$

est bijective. (Elle est injective : $(12)\, \alpha = (12)\, \beta \Rightarrow \alpha = \beta$; et bijective par suite du théorème 3 (chap. 1, § 5). On peut le démontrer également en remarquant que $(\rho_{(12)})^2$ est une application identique.) Puisque $\varepsilon_{(12)\pi} = \varepsilon_{(12)} \varepsilon_\pi = -\varepsilon_\pi$, on a $\rho_{(12)} A_n = \bar{A}_n$, $\rho_{(12)} \bar{A}_n = A_n$. Par conséquent, le nombre de permutations paires de S_n coïncide avec celui de permutations impaires, d'où $|A_n| = \frac{1}{2} |S_n| = \frac{n!}{2}$. ■

COROLLAIRE 2. — *Soit* $\pi \in S_n$ *une permutation décomposée en un produit de cycles indépendants de longueurs* $l_1, l_2, \ldots, l_m$. *Alors,*

$$\varepsilon_\pi = (-1)^{\sum_{k=1}^{m} (l_k - 1)}.$$

En effet, d'après le théorème 5 on a $\varepsilon_\pi = \varepsilon_{\pi_1 \ldots \pi_m} = \varepsilon_{\pi_1} \ldots \varepsilon_{\pi_m}$. En outre, $\varepsilon_{\pi_k} = (-1^{l_k - 1})$ car π_k s'écrit sous la forme d'un

produit de $l_k - 1$ transpositions (voir démonstration du corollaire 2 du théorème 4). Finalement, on a

$$\varepsilon_\pi = (-1)^{l_1-1} \dots (-1)^{l_m-1} = (-1)^{\sum\limits_{k=1}^{m}(l_k-1)} . \blacksquare$$

Pour nous délasser un peu après l'étude des choses sérieuses, considérons en conclusion un petit jeu bien connu. Quinze fiches plates carrées de même dimension, numérotées de 1 à 15, sont placées sur un tableau carré divisé en 16 cases ayant mêmes dimensions que les fiches. Une seule case reste libre ; en l'utilisant on peut déplacer les fiches horizontalement et verticalement (sans les enlever du tableau). Etant donné une disposition arbitraire des fiches

i_1	i_2	i_3	i_4
i_5	i_6	i_7	i_8
i_9	i_{10}	i_{11}	i_{12}
i_{13}	i_{14}	i_{15}	

a)

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	

b)

Fig. 12.

(fig. 12, *a*) on demande de passer à leur disposition correcte dans l'ordre croissant des numéros (fig. 12, *b*) ; pour la case libre à l'état de départ on peut prendre le coin inférieur droit. Quand un tel passage est-il possible ? On trouve la réponse dans la théorie élémentaire des groupes qui « tue » le jeu lorsqu'il est en vogue dans les salons. On associe aux figures *a*) et *b*) une permutation $\pi \in S_{15}$ et on s'assure sans peine (nous conseillons au lecteur de le faire quand même) que la disposition correcte peut être obtenue si, et seulement si, la parité ε_π de la permutation π est égale à 1, c'est-à-dire si $\pi \in A_{15}$.

EXERCICES

1. Montrer que si $M = \langle S \rangle$ est un monoïde engendré par un ensemble S et si tout élément $s \in S$ est inversible dans M, alors M est un groupe.

2. Le groupe est un monoïde G dans lequel les équations de la forme $ax = b$, $ya = b$ sont résolubles de façon unique, quels que soient $a, b \in G$. Démontrer cette assertion.

3. Montrer que l'ensemble $A_1(\mathbb{R})$ des *applications* dites *affines* $\varphi_{a,b} : x \mapsto ax + b$ $(a, b \in \mathbb{R}; a \neq 0)$ de la droite réelle $\mathbb{R}$ muni de la loi de composition des applications $\varphi_{a,b}\varphi_{c,d} = \varphi_{ac,ad+b}$ a une structure de groupe. Le groupe $A_1(\mathbb{R})$ contient le sous-groupe $\mathrm{GL}(1, \mathbb{R})$ qui laisse fixe le point $x = 0$ et le sous-groupe des « translations pures » $x \mapsto x + b$.

4. Le groupe $\mathrm{SL}(2, \mathbb{Z})$ contient les éléments $A = \begin{Vmatrix} 0 & 1 \\ -1 & 0 \end{Vmatrix}$, $B = \begin{Vmatrix} 0 & 1 \\ -1 & -1 \end{Vmatrix}$ d'ordres 4 et 3 respectivement. Montrer que $\langle AB \rangle$ est un sous-groupe cyclique infini de $\mathrm{SL}(2, \mathbb{Z})$. Ainsi, le produit de deux éléments d'ordre fini dans le groupe G ne doit pas nécessairement être un élément d'ordre fini. En est-il de même dans un groupe abélien ?

5. Démontrer que le groupe G d'ordre pair $|G| = 2n$ contient nécessairement un élément g d'ordre 2. (I n d i c a t i o n. Considérer la partition de G en couples g, g^{-1}.)

6. Démontrer que $S_n = \langle(12), (13), \ldots, (1n)\rangle$.

7. Démontrer que $S_n = \langle(12), (123 \ldots n)\rangle$.

8. Démontrer que le groupe alterné A_n, $n \geqslant 3$, est engendré par les cycles de longueur 3, et qu'on a en effet

$$A_n = \langle(123), (124), \ldots, (12n)\rangle.$$

9. Déterminer la signature de la permutation

$$\pi = \begin{pmatrix} 1 & 1 & 3 & \ldots & n-1 & n \\ n & n-1 & n-2 & \ldots & 2 & 1 \end{pmatrix}.$$

10. Soit $\Omega \times \Omega$ le carré cartésien de $\Omega = \{1, 2, \ldots, n\}$. Un couple $(i, j) \in \in \Omega \times \Omega$ est appelé *inversion par rapport à la permutation* $\sigma \in S_n$ (ou tout court σ-*inversion*) si $i < j$ mais $\sigma(i) > \sigma(j)$. Posons

$$\operatorname{sgn} \sigma = \prod_{1 \leqslant i < j \leqslant n} \frac{\sigma(j) - \sigma(i)}{j - i}.$$

Puisque $(\sigma(j) - \sigma(i))/(j - i)$ est un nombre rationnel non nul qui est négatif si, et seulement si, (i, j) est une σ-inversion, et comme $\sigma\colon \Omega \to \Omega$ est une application bijective, on a $\operatorname{sgn} \sigma = (-1)^k$, où k est le nombre total de σ-inversions. Si $\tau = (ij)$ est une transposition, alors $\operatorname{sgn} \tau = -1$. Il est facile de voir que

$$(\sigma(j)\,\sigma(i))\,\sigma = \begin{pmatrix} \ldots\sigma(j)\ldots\sigma(i)\ldots \\ \ldots\sigma(i)\ldots\sigma(j)\ldots \end{pmatrix} \begin{pmatrix} \ldots i & \ldots j & \ldots \\ \ldots\sigma(i)\ldots\sigma(j)\ldots \end{pmatrix} = \begin{pmatrix} \ldots i & \ldots j & \ldots \\ \ldots\sigma(j)\ldots\sigma(i)\ldots \end{pmatrix},$$

de sorte que la σ-inversion (i, j) cesse d'être une inversion par rapport à la permutation $\tau\sigma$, où $\tau = (\sigma(j)\,\sigma(i))$ est une transposition. Montrer qu'il existe k transpositions $\tau_1, \ldots, \tau_k$ telles que $\tau_k\tau_{k-1} \ldots \tau_1\sigma = e$ est une permutation identique. Donc, $\sigma = \tau_1 \ldots \tau_{k-1}\tau_k$ et $\operatorname{sgn} \sigma = (-1)^k = \varepsilon_\sigma$ sont deux notations équivalentes d'un seul et même invariant de la permutation (la notation sgn provient du mot latin *signum*). Nous avons obtenu encore un procédé commode pour déterminer le signe d'une permutation. Par exemple, l'ensemble des inversions par rapport à la permutation (4) se compose de cinq couples (1, 5) (2, 5), (3, 5), (4, 5), (6, 7), si bien que $\operatorname{sgn} \pi = -1$. Pratiquement le problème se ramène au calcul, dans la ligne inférieure de la permutation π, du nombre des entiers j supérieurs à i, mais placés devant i, quand $i = 1, 2, \ldots, n - 1$.

11. Démontrer qu'un sous-ensemble non vide H d'un groupe fini (multiplicatif) G est un sous-groupe si H est stable pour la multiplication. Cela signifie que dans ce cas les exigences d'existence dans H d'un élément unité e et d'un inverse h^{-1} pour tout $h \in H$ sont superflues.

12. Quel système de générateurs peut-on proposer pour le groupe multiplicatif $(\mathbb{Q}_+, \cdot)$ des nombres rationnels positifs? (I n d i c a t i o n. Utiliser le théorème fondamental de l'arithmétique (chap. 1, § 8).) Un système fini de générateurs existe-t-il dans $(\mathbb{Q}_+, \cdot)$?

13. Démontrer que la puissance k-ième π^k du cycle $\pi = (12 \ldots n) \in S_n$ est un produit de $d = \text{P.G.C.D.}(n, k)$ cycles indépendants dont chacun a la longueur $q = n/d$.

14. Soient $A, B \in M_n(\mathbb{R})$ et $(AB)^m = E$ pour un entier m. Est-il vrai que $(BA)^m = E$?

10—0877

§ 3. Morphismes des groupes

1. Isomorphismes.— Comme nous l'avons indiqué plus haut trois rotations φ_0, φ_1, φ_2 en sens inverse des aiguilles d'une montre, d'angles 0°, 120°, 240°, transforment un triangle équilatéral P_3 en lui-même. Il existe encore trois *transformations de symétrie axiale* (*réflexions*) ψ_1, ψ_2, ψ_3 avec les axes de symétrie 1-1′, 2-2′, 3-3′ indiqués

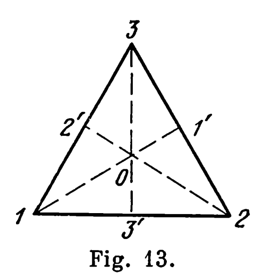


Fig. 13.

sur la fig. 13. A toutes les six transformations de symétrie il correspond des permutations sur l'ensemble des sommets du triangle. On obtient

$$\varphi_0 \sim e, \quad \varphi_1 \sim (123), \quad \varphi_2 \sim (132),$$

$$\psi_1 \sim (23), \quad \psi_2 \sim (13), \quad \psi_3 \sim (12).$$

Puisqu'il n'existe pas d'autres permutations de degré 3, on peut affirmer que le groupe D_3 de toutes les transformations de symétrie d'un triangle équilatéral présente une forte ressemblance avec le groupe symétrique S_3.

De ce point de vue, ressemblent l'un à l'autre les groupes cycliques C_n (voir exemple du § 2, n° 3) et $\langle(12\ \dots\ n)\rangle \subset S_n$. Ces faits, ainsi que des réflexions générales sur les groupes, obligent à poser une question bien naturelle relative aux propriétés fondamentales des groupes. Il semble à première vue qu'une information complète soit contenue dans la table de multiplication du groupe G appelé *table de Cayley*:

	g_1	g_2	$\dots$	g_n	$\dots$
g_1	g_1g_1	g_1g_2	$\dots$	g_1g_n	$\dots$
g_2	g_2g_1	g_2g_2	$\dots$	g_2g_n	$\dots$
$\cdot$					
$\cdot$	$\cdot\ \cdot\ \cdot$	$\cdot\ \cdot\ \cdot$	$\cdot\ \cdot\ \cdot$	$\cdot\ \cdot\ \cdot$	$\cdot\ \cdot\ \cdot$
$\cdot$					
g_n	g_ng_1	g_ng_2	$\dots$	g_ng_n	$\dots$
$\cdot$					
$\cdot$	$\cdot\ \cdot\ \cdot$	$\cdot\ \cdot\ \cdot$	$\cdot\ \cdot\ \cdot$	$\cdot\ \cdot\ \cdot$	$\cdot\ \cdot\ \cdot$
$\cdot$					

En effet, plusieurs propriétés du groupe peuvent être déduites en analysant la table de Cayley ou, ce qui revient au même, la matrice $M = (m_{ij})$ (de type (n, n) si $n = (G:e)$) d'éléments $m_{ij} = g_i g_j \in G$. Nous remarquons, par exemple, que dans chaque ligne et dans chaque colonne de la matrice M tout élément du groupe G ne se rencontre qu'une seule fois (voir plus loin la démonstration du théorème 2). Le groupe G est commutatif si, et seulement si, la matrice M est symétrique, c'est-à-dire si $m_{ij} = m_{ji}$. On pourrait continuer cette liste de propriétés, mais il est tout de même assez difficile de comparer deux tables pour les groupes G et G' de même ordre, car la forme de la matrice M dépend de la numération (de la disposition) des éléments du groupe. Dans le cas des groupes infinis la situation se complique davantage.

L'approche la plus correcte et la plus radicale du problème de distinction (ou au contraire d'identification) des groupes G et G' est proposée par la notion d'isomorphisme.

DÉFINITION. — *Deux groupes G et G' munis des opérations $*$ et $\circ$ sont dits isomorphes s'il existe une application $f: G \to G'$ telle que*
(i) *$f(a * b) = f(a) \circ f(b)$ pour tout couple $a, b \in G$;*
(ii) *f est bijective.*

L'isomorphisme des groupes est souvent noté $G \cong G'$.

Indiquons les propriétés les plus simples de l'isomorphisme.

1) *L'élément neutre se transforme en élément neutre.* En effet, si e est l'élément neutre du groupe G, on a $e * a = a * e = a$ et donc $f(e) \circ f(a) = f(a) \circ f(e) = f(a)$, d'où il résulte que $f(e) = e'$ est l'élément neutre du groupe G'. Ce raisonnement utilise, bien qu'en partie, les deux propriétés de f. Pour (i) cela est évident, alors que la propriété (ii) assure la surjectivité de f, de sorte que le groupe G' tout entier ne contient que les éléments $f(g)$. ■

2) $f(a^{-1}) = f(a)^{-1}$. En effet, d'après 1), $f(a) \circ f(a^{-1}) = f(a * a^{-1}) = f(e) = e'$ est l'élément neutre de G', d'où

$$f(a)^{-1} = f(a)^{-1} \circ e' = f(a)^{-1} \circ (f(a) \circ f(a^{-1})) =$$
$$= (f(a)^{-1} \circ f(a)) \circ f(a^{-1}) = e' \circ f(a^{-1}) = f(a^{-1}). \quad ■$$

3) *L'application réciproque $f^{-1}: G' \to G$ (qui existe en vertu de la propriété* (ii)) *est aussi un isomorphisme.*

En raison du corollaire au théorème 2 (chap. 1, § 5) il suffit de s'assurer que f^{-1} vérifie la propriété (i). Soient $a', b' \in G'$. Alors, f étant bijective, on a $a' = f(a)$, $b' = f(b)$ pour certains $a, b \in G$. Puisque f est un isomorphisme, $a' \circ b' = f(a) \circ f(b) = f(a * b)$. D'où l'on tire $a * b = f^{-1}(a' \circ b')$. Etant donné $a = f^{-1}(a')$, $b = f^{-1}(b')$, on a $f^{-1}(a' \circ b') = f^{-1}(a') * f^{-1}(b')$. ■

Une vérification assez simple montre que la correspondance $\sim$ entre les groupes D_3 et S_3 que nous avons établie, est en effet un isomorphisme.

A titre de l'application isomorphe f du groupe multiplicatif $(\mathbb{R}_+, \cdot)$ des nombres réels positifs sur le groupe additif $(\mathbb{R}, +)$ de tous les nombres réels, on peut indiquer $f = \ln$. La propriété connue du logarithme $\ln ab = \ln a + \ln b$ interprète justement la propriété (i) dans la définition de l'isomorphisme. L'application réciproque de f est $x \mapsto e^x$.

Démontrons maintenant deux théorèmes généraux qui mettent en évidence le rôle de l'isomorphisme dans la théorie des groupes.

THÉORÈME 1. — *Tous les groupes cycliques de même ordre (y compris d'ordre infini) sont isomorphes.*

DÉMONSTRATION. — En effet, si $\langle g \rangle$ est un groupe cyclique infini, toutes les puissances g^n de l'élément générateur g sont distinctes, et nous obtenons un isomorphisme $f: \langle g \rangle \to (\mathbb{Z}, +)$, en posant $g^n \mapsto f(g^n) = n$. Il est évident que f est bijective, quant à la propriété $f(g^m g^n) = f(g^n) + f(g^m)$, elle découle du théorème 2 du § 2.

Soient maintenant $G = \{e, g, \ldots, g^{q-1}\}$ et $G' = \{e', g', \ldots, (g')^{q-1}\}$ deux groupes cycliques d'ordre q (nous ne distinguons pas les opérations dans G et G'). Définissons l'application bijective

$$f: g^k \mapsto (g')^k, \quad k = 0, 1, \ldots, q-1.$$

En posant $n + m = lq + r$, $0 \leqslant r \leqslant q-1$ quels que soient n, $m = 0, 1, \ldots, q-1$ et en raisonnant comme lors de la démonstration du théorème 3 du § 2, on aura

$$f(g^{n+m}) = f(g^r) = (g')^r = (g')^{n+m} = (g')^n (g')^m = f(g^n) f(g^m). \blacksquare$$

THÉORÈME 2 (Cayley). — *Tout groupe fini d'ordre n est isomorphe à un sous-groupe d'un groupe symétrique S_n.*

DÉMONSTRATION. — Soit G notre groupe, $n = |G|$. On peut poser que S_n est le groupe de toutes les applications bijectives de l'ensemble G sur lui-même, la nature des éléments permutés par les éléments de S_n étant sans importance.

Etant donné un élément $a \in G$ considérons l'application $L_a : G \to G$ définie par la formule

$$L_a(g) = ag.$$

Si $e = g_1, g_2, \ldots, g_n$ sont tous les éléments du groupe G, alors a, $ag_2, \ldots, ag_n$ seront les mêmes éléments, mais disposés dans un autre ordre (rappelons-nous la table de Cayley!). Cela se comprend, car

$$ag_i = ag_j \Rightarrow a^{-1}(ag_i) = a^{-1}(ag_j) \Rightarrow (a^{-1}a) g_i = (a^{-1}a) g_j \Rightarrow g_i = g_j.$$

Par conséquent, L_a est une application bijective (une permutation) dont l'inverse sera $L_a^{-1} = L_{a^{-1}}$. L'application identique est naturellement L_e.

En utilisant de nouveau l'associativité de la multiplication dans G, on obtient $L_{ab}(g) = (ab)\,g = a\,(bg) = L_a\,(L_b g)$, c'est-à-dire $L_{ab} = L_a \circ L_b$.

Ainsi, l'ensemble $L_e, L_{g_2}, \ldots, L_{g_n}$ forme un sous-groupe, disons H, du groupe $S(G)$ de toutes les bijections de l'ensemble G sur lui-même, c'est-à-dire du groupe S_n. On a l'inclusion $H \subset S_n$ et la correspondance $L: a \mapsto L_a \in H$ qui vérifie, en vertu de ce qui précède, toutes les propriétés de l'isomorphisme. ■

Malgré sa simplicité, le théorème de Cayley revêt une grande importance dans la théorie des groupes. Il dégage un certain être universel (la famille $\{S_n \mid n = 1, 2, \ldots\}$ de groupes symétriques) qui contient en général tous les groupes finis considérés à un isomorphisme près. L'expression « à un isomorphisme près » reflète l'essence non seulement de la théorie des groupes qui cherche à réunir en une seule classe tous les groupes isomorphes, mais aussi de toutes les mathématiques qui seraient dénuées de sens si elles ne procédaient pas à de telles généralisations.

En posant $G' = G$ dans la définition de l'isomorphisme, nous obtenons une application isomorphe $\varphi: G \to G$ du groupe G sur lui-même. Elle s'appelle *automorphisme* du groupe G. Par exemple, l'application identique $e_G: g \mapsto g$ (que nous désignerons dans la suite tout simplement par 1) est un automorphisme. En général, G possède encore des automorphismes non triviaux. La propriété 3) des applications isomorphes montre qu'une application réciproque d'un automorphisme est aussi un automorphisme. Si φ, ψ sont des automorphismes du groupe G, on a $(\varphi \circ \psi)\,(ab) = \varphi\,(\psi\,(ab)) = \varphi\,(\psi\,(a)\,\psi\,(b)) = (\varphi \circ \psi)\,(a) \cdot (\varphi \circ \psi)\,(b)$ quels que soient $a, b \in G$. Cela signifie que l'ensemble Aut (G) de tous les automorphismes du groupe G forme un groupe qui est un sous-groupe du groupe $S(G)$ de toutes les applications bijectives $G \to G$.

2. Homomorphismes.— Le groupe des automorphismes Aut (G) du groupe G contient un sous-groupe particulier. Ce dernier est désigné par le symbole Inn (G) et appelé *groupe des automorphismes internes.* Ses éléments sont les applications

$$I_a: g \mapsto aga^{-1}$$

Un petit exercice montre que I_a vérifie réellement toutes les propriétés que doit posséder un automorphisme et que $I_a^{-1} = I_{a^{-1}}$, $I_e = 1$ est un automorphisme identique, $I_a \circ I_b = I_{ab}$ (parce que $(I_a \circ I_b)\,(g) = I_a\,(I_b\,(g)) = I_a\,(bgb^{-1}) = abgb^{-1}a^{-1} = abg\,(ab)^{-1} = I_{ab}\,(g)$).

La dernière relation exprime que l'application

$$f : G \to \text{Inn}(G)$$

du groupe G dans le groupe Inn (G) de ses automorphismes internes, définie par la formule $f(a) = I_a$, $a \in G$, possède la propriété (i) de l'application isomorphe: $f(a) \circ f(b) = f(ab)$. Pourtant, dans ces conditions, la propriété (ii) ne doit pas être nécessairement vérifiée. Si par exemple G est un groupe abélien, on a $aga^{-1} = g$ pour tous les $a, g \in G$ si bien que $I_a = I_e$ et le groupe Inn (G) est constitué d'un seul élément unité I_e. Cette circonstance rend naturelle la définition générale suivante:

DÉFINITION. — *L'application $f: G \to G'$ d'un groupe $(G, *)$ dans $(G', \circ)$ est un homomorphisme, si*

$$f(a * b) = f(a) \circ f(b), \quad \forall a, b \in G$$

(autrement dit, la propriété (ii) figurant dans la définition de l'isomorphisme est omise).

On appelle noyau de l'homomorphisme f l'ensemble

$$\text{Ker}\, f = \{g \in G \mid f(g) = e', \text{ élément unité du groupe } G'\}.$$

L'application homomorphe d'un groupe dans lui-même est encore appelée *endomorphisme.*

Cette définition n'exige pas non seulement que f soit bijective mais aussi qu'elle soit surjective (c'est-à-dire une application « sur ») ce qui d'ailleurs importe peu, car on peut toujours se contenter de considérer l'image $\text{Im}\, f \subset G'$ qui est évidemment un sous-groupe de G'. La différence principale que l'homomorphisme f présente par rapport à l'isomorphisme, est l'existence d'un noyau non trivial $\text{Ker}\, f$ qui constitue en quelque sorte une mesure de non-injectivité de f. Si $\text{Ker}\, f = \{e\}$, alors $f: G \to \text{Im}\, f$ est un isomorphisme.

Remarquons que

$$f(a) = e', \ f(b) = e' \Rightarrow f(a * b) = f(a) \circ f(b) = e' \circ e' = e'$$

et

$$f(a^{-1}) = f(a)^{-1} = (e')^{-1} = e'.$$

C'est pourquoi le noyau $\text{Ker}\, f$ est un sous-groupe de G. Soit $H = \text{Ker}\, f \subset G$. Alors (nous omettons maintenant les signes $*$ et $\circ$):

$$f(ghg^{-1}) = f(g)\, f(h)\, f(g)^{-1} = f(g)\, e' f(g)^{-1} = e', \quad \forall\, h \in H, \ g \in G,$$

c'est-à-dire $ghg^{-1} \in H$ et donc $gHg^{-1} \subset H$. En remplaçant ici g par g^{-1}, on obtient $g^{-1}Hg \subset H$, d'où $H \subset gHg^{-1}$. Par suite,

$$gHg^{-1} = H, \quad \forall g \in G.$$

Les sous-groupes vérifiant cette propriété sont appelés *sous-groupes distingués* (on les appelle encore *sous-groupes invariants* ou *normaux*). Ainsi, nous avons démontré le théorème suivant :

THÉORÈME 3. — *Les noyaux des homomorphismes sont toujours des sous-groupes distingués.* ∎

C'est beaucoup plus tard que nous pourrons apprécier pleinement l'importance de ce fait. Pour l'instant, remarquons que tout sous-groupe de G est loin d'être distingué. Par exemple, dans S_3, le sous-groupe cyclique $\langle(123)\rangle = A_3$ est distingué, alors que $\langle(12)\rangle =$ $= \{e, (12)\}$ ne l'est pas (il n'est pas recommandé d'appeler $\langle(12)\rangle$ « un sous-groupe non distingué »).

3. Terminologie. Exemples. — Il y a lieu de signaler que les termes de surjection, d'injection et de bijection employés relativement aux applications des ensembles quels qu'ils soient (non munis de lois de composition) sont remplacés dans le cas des groupes (et d'autres structures algébriques) respectivement par les termes d'épimorphisme (homomorphisme « sur »), de monomorphisme (homomorphisme à noyau unité) et d'isomorphisme (homomorphisme bijectif, c'est-à-dire épimorphisme et monomorphisme à la fois). La tendance actuelle est au remplacement de l'homomorphisme par le terme morphisme. Il est utile d'avoir en vue ces conventions terminologiques lors de la lecture des ouvrages mathématiques, mais au début on peut se contenter de deux termes : isomorphisme et homomorphisme avec adjonction de « dans » ou « sur ».

Nous donnons ci-dessous encore quelques exemples de morphismes des groupes.

1) L'application du groupe additif des entiers relatifs $\mathbb{Z}$ sur un groupe cyclique fini $\langle g\rangle$ d'ordre q est un homomorphisme si l'on pose $f: n \mapsto g^n$ (voir théorème 2 du § 2). Dans ce cas on a évidemment $\operatorname{Ker} f = \{lq \mid l \in \mathbb{Z}\}$. En effet, il est clair que $\{lq\} \subset \operatorname{Ker} f$. L'inclusion réciproque découle du théorème 3 du § 2.

2) L'application $f: \mathbb{R} \to T = \mathrm{SO}(2)$ du groupe additif des nombres réels sur le groupe T des rotations du plan à point fixe O, définie par la formule $f(\lambda) =$ $= \Phi_\lambda$ (Φ_λ est une rotation de $2\pi\lambda$ en sens inverse des aiguilles d'une montre) est un homomorphisme, puisque $\Phi_\lambda \circ \Phi_\mu = \Phi_{\lambda+\mu}$. La rotation d un angle égal à un multiple entier de 2π, coïncidant avec la rotation identique (d'angle zéro), on a $\operatorname{Ker} f = \{2\pi n \mid n \in \mathbb{Z}\}$. On dit aussi que f est un homomorphisme de $\mathbb{R}$ sur une circonférence S^1 de rayon unité, car il existe une bijection entre les Φ_λ et les points de coordonnées polaires $(1, 2\pi\lambda)$, $0 \leqslant \lambda < 1$, sur S^1.

3) L'application du groupe linéaire complet $\mathrm{GL}(n)$ des matrices A à coefficients réels (c'est-à-dire dans $\mathbb{R}$) de déterminant $\det A$ non nul, sur le groupe multiplicatif $\mathbb{R}^*$ des nombres réels non nuls est un homomorphisme si l'on pose $f = \det$. La condition d homomorphisme $f(AB) = f(A)\, f(B)$ n'est qu'un autre énoncé du théorème 5 (chap. 3, § 2). Par définition $\mathrm{SL}(n) = \operatorname{Ker} f$.

4) Considérons le groupe cyclique $C_2 = \langle -1 \rangle = \{1, -1\}$ d'ordre 2. Si l'on veut, on peut le définir de façon abstraite par la table de Cayley:

$$C_2: \begin{array}{r|rr} & 1 & -1 \\ \hline 1 & 1 & -1 \\ -1 & -1 & 1 \end{array}$$

L'application $S_n \to C_2$ à l'aide de la fonction connue $\varepsilon = \text{sgn}: \pi \mapsto \varepsilon_\pi$ (la signature de la permutation π) est un homomorphisme du groupe symétrique S_n sur C_2. Ceci étant, on a selon la définition du groupe alterné, $\text{Ker}\ \varepsilon = A_n$.

5) Un groupe infini peut être isomorphe à son sous-groupe propre. En effet, le groupe additif $(\mathbb{Z}, +)$ contient le sous-groupe propre $n\mathbb{Z} = \{nk \mid k \in \mathbb{Z}\}$, où $n > 1$ est un entier naturel fixe. On vérifie sans peine que l'application $g_n: \mathbb{Z} \to n\mathbb{Z}$ définie par la relation $g_n(k) = nk$ est un isomorphisme. Remarquons en passant que $\mathbb{Z}$ et $n\mathbb{Z}$ sont des groupes cycliques infinis dont les éléments générateurs sont respectivement 1 ou -1 et n ou $-n$; c'est pourquoi g_n et l application $k \mapsto -nk$ sont les seuls isomorphismes $\mathbb{Z} \to n\mathbb{Z}$.

6) Le groupe Aut (G) et même n'importe quel élément $\varphi \in$ Aut (G) différent de l'élément unité peuvent fournir beaucoup de renseignements importants sur le groupe G. Voilà un exemple suggestif. Soit G un groupe fini auquel on applique un automorphisme φ d'ordre 2 ($\varphi^2 = 1$) sans points fixes:

$$a \neq e \Rightarrow \varphi(a) \neq a.$$

Supposons que $\varphi(a)a^{-1} = \varphi(b)b^{-1}$ pour certains $a, b \in G$. La multiplication de cette égalité à gauche par $\varphi(b)^{-1}$ et à droite par a donne alors $\varphi(b)^{-1}\varphi(a) = b^{-1}a$, c'est-à-dire $\varphi(b^{-1}a) = b^{-1}a$, d'où $b^{-1}a = e$ et $b = a$. Ainsi, $\varphi(a)a^{-1}$ parcourt avec a les valeurs de tous les éléments du groupe G ou, ce qui revient au même, tout élément $g \in G$ s'écrit sous la forme de $g = \varphi(a)a^{-1}$. Or, dans un tel cas $\varphi(g) = \varphi(\varphi(a))\varphi(a^{-1}) = \varphi^2(a)\varphi(a^{-1}) = a\varphi(a)^{-1} = (\varphi(a)a^{-1})^{-1} = g^{-1}$. Ainsi, φ coïncide avec l'application $g \mapsto g^{-1}$. Connaissant ce fait, on obtient $ab = \varphi(a^{-1})\varphi(b^{-1}) = \varphi(a^{-1}b^{-1}) = (a^{-1}b^{-1})^{-1} = ba$, c'est-à-dire le groupe G est un groupe abélien! De plus, $(G: e)$ est un nombre impair, car G est constitué de e et de couples disjoints d'éléments g_i, $g_i^{-1} = \varphi(g_i)$.

7) Considérons un exemple (voir aussi § 1, exercice 3) qui montre dans quelle mesure on peut modifier l'opération définie sur un groupe sans changer, au sens de l'isomorphisme, le groupe lui-même. Soit G un groupe arbitraire et soit t son élément fixe quelconque. Définissons sur l'ensemble G une nouvelle opération:

$$(g, h) \mapsto g * h = gth.$$

On vérifie immédiatement que $(g_1 * g_2) * g_3 = g_1 * (g_2 * g_3)$, c'est-à-dire que l'opération $*$ est associative. En outre $g * t^{-1} = t^{-1} * g = g$ et $g * (t^{-1}g^{-1}t^{-1}) = (t^{-1}g^{-1}t^{-1}) * g = t^{-1}$, ce qui signifie que $(G, *)$ est un groupe possédant un élément unité $e_* = t^{-1}$. L élément inverse de g dans $(G, *)$ est $g_*^{-1} = t^{-1}g^{-1}t^{-1}$. L'application $f: g \mapsto gt^{-1}$ établit un isomorphisme entre les groupes $(G, \cdot)$ et $(G, *)$, c'est-à-dire $f(gh) = f(g) * f(h)$.

Notons, entre autres, que tous les exemples donnés ci-dessus illustrent une règle générale: l'étude des morphismes du groupe G permet d'obtenir une information importante sur le groupe lui-même.

4. Classes suivant un sous-groupe. La définition de l'homomorphisme $f: G \to G'$ et les exemples considérés plus haut entraînent qu'à tous les éléments de l'ensemble

$$a\ \text{Ker}\ f = \{ab \mid b \in \text{Ker}\ f\}, \quad a \in G,$$

est associé un seul et même élément $f(a)$ du groupe G' : $f(ab) = f(a) f(b) = f(a) e' = f(a)$. Réciproquement, si $f(g) = f(a)$, on a $f(a^{-1}g) = f(a^{-1}) f(g) = f(a)^{-1} f(g) = e'$, d'où $a^{-1}g = b \in \operatorname{Ker} f$ et $g = ab \in a \operatorname{Ker} f$. Ce fait montre qu'il y a intérêt à partager G en sous-ensembles de la forme $a \operatorname{Ker} f$. Proposons-nous d'étudier une telle partition dans le cas général, indépendamment des homomorphismes.

DÉFINITION. — *Soit H un sous-groupe d'un groupe G. On appelle classe à gauche du groupe G suivant le sous-groupe H (ou tout simplement G suivant H) l'ensemble gH des éléments gh, où g est un élément fixe de G, et h parcourt H. L'élément g s'appelle représentant de la classe gH.*

On définit de manière analogue les *classes à droite Hg*. Parfois, les classes à gauche définies comme ci-dessus sont appelées classes à droite et vice versa. Il importe, donc, d'adopter l'une ou l'autre des terminologies. Si $H = \operatorname{Ker} f$ est le noyau de l'homomorphisme, alors $gH = Hg$ car le sous-groupe H de G est distingué (voir n° 2). Remarquons que l'une des classes est constituée par le sous-groupe H lui-même, $H = He = eH$. Aucune autre classe n'est un sous-groupe. En effet, si gH était un sous-groupe, on aurait $e \in gH$, d'où $e = gh$, $g = h^{-1}$, et $gH = h^{-1}H = H$.

THÉORÈME 4.— *Deux classes à gauche de G suivant H coïncident ou ne possèdent aucun élément commun. La partition de G en classes à gauche suivant H définit sur G une relation d'équivalence.*

DÉMONSTRATION. — Supposons que les classes g_1H et g_2H possèdent un élément commun $a = g_1h_1 = g_2h_2$. Alors, $g_2 = g_1h_1h_2^{-1}$, et tout élément g_2h de la classe g_2H est de la forme $g_1h_1h_2^{-1}h = g_1h'$, où $h' = h_1h_2^{-1}h \in H$. Par conséquent, $g_2H \subset g_1H$. On démontre de même que tout élément de la classe g_1H est contenu dans g_2H, et donc $g_1H = g_2H$.

Puisque tout élément $g \in G$ donné d'avance est contenu dans gH, le raisonnement développé montre que G se présente comme réunion des classes à gauche disjointes suivant le sous-groupe H:

$$G = \bigcup g_iH.$$

En vertu du principe général exposé au chapitre 1, § 6, cette partition induit sur G une relation d'équivalence définie de façon suivante:

$$a \sim b \Leftrightarrow a^{-1}b \in H.$$

Si l'on veut, on peut s'assurer immédiatement que cette relation est réflexive, symétrique et transitive: $a \sim a$, car $a^{-1}a = e \in H$; $a \sim b \Leftrightarrow a^{-1}b = h \Leftrightarrow b^{-1}a = h^{-1} \in H \Leftrightarrow b \sim a$; $a \sim b$, $b \sim c \Rightarrow b^{-1}a = h_1$, $c^{-1}b = h_2 \Rightarrow c^{-1}a = c^{-1}bh_1 = h_2h_1 \in H \Rightarrow a \sim c$. ■

Une assertion analogue est vraie pour les classes à droite.

La partition en classes suivant un sous-groupe apparaît de façon naturelle dans les groupes de permutations. Soit, par exemple, $G = S_n$ le groupe symétrique opérant sur l'ensemble $\Omega = \{1, 2, \ldots, n\}$. Si l'on considère l'ensemble H d'éléments $\pi \in S_n$ tels que $\pi(n) = n$, il n'est pas difficile de s'assurer que H est un sous-groupe de S_n qui peut être identifié avec S_{n-1}. Soient $\tau_0 = e$ et $\tau_i = (in)$ une transposition qui transforme n en i ($i = 1, 2, \ldots, n-1$). Il est clair que

$$S_n = \bigcup_{k=0}^{n-1} \tau_k S_{n-1}.$$

Considérons la partition de S_3 en classes à gauche et à droite suivant le sous-groupe $\langle(12)\rangle = S_2$:

$$S_3 = \{e, (12)\} \cup \{(13), (123)\} \cup \{(23), (132)\};$$

$$S_3 = \{e, (12)\} \cup \{(13), (132)\} \cup \{(23), (123)\}.$$

On voit que l'ensemble des classes à gauche gS_2 ne coïncide pas avec l'ensemble des classes à droite S_2g'. Néanmoins entre les ensembles $\{gH\}$ et $\{Hg'\}$ il existe toujours une correspondance biunivoque car

$$x = gh \in gH \leftrightarrow x^{-1} = h^{-1}g^{-1} \in Hg^{-1}.$$

En effet, si par exemple $h_1g_1^{-1} = h_2g_2^{-1}$, alors $g_1 = g_2h_2^{-1}h_1$ et $g_1H = g_2H$. En particulier, *si* $\{e, x, y, z, \ldots\}$ *est l'ensemble des représentants des classes à gauche* (*respectivement à droite*), *alors* $\{e, x^{-1}, y^{-1}, z^{-1}, \ldots\}$ *est l'ensemble des représentants des classes à droite* (*respectivement à gauche*). *Les puissances de ces ensembles coïncident.* ∎

On convient de désigner l'ensemble de toutes les classes à gauche de G suivant H par le symbole G/H (ou $(G/H)_l$ s'il s'avère nécessaire de considérer à la fois l'ensemble $(G/H)_r$ des classes à droite de G suivant H). Pour la puissance Card G/H de cet ensemble on emploie l'appellation « indice du sous-groupe H dans G » et on introduit une désignation spéciale $(G: H)$ qui s'accorde bien avec la désignation $(G: e)$ de l'ordre $|G|$ du groupe G (nombre de classes suivant un sous-groupe unitaire). L'application $H \to gH$ étant biunivoque (rappelons-nous la démonstration du théorème de Cayley et l'application L_g), Card $gH = (H: e)$. Ainsi, on obtient une formule facile à retenir

$$(G:e) = (G:H)\ (H:e)$$

dont il ressort le théorème classique suivant:

THÉORÈME 5 (Lagrange). — *L'ordre d'un groupe fini est divisible par l'ordre de chacun de ses sous-groupes.* ∎

COROLLAIRE. — *L'ordre d'un élément divise l'ordre du groupe. Le groupe d'ordre p premier est toujours un groupe cyclique qui est unique à un isomorphisme près.*

En effet, l'ordre de tout élément $g \in G$ coïncide avec celui du sous-groupe cyclique $\langle g \rangle$ engendré par $\langle g \rangle$ (théorème 3 du § 2). Si $|G| = p$ est un nombre premier, et H est un sous-groupe non unitaire, la divisibilité de p par $|H|$ signifie que $|H| = p$, d'où $H = G$. Donc, G coïncide avec le sous-groupe cyclique engendré par un élément $g \neq e$. Tous les groupes cycliques d'ordre donné étant isomorphes (théorème 1), on a le droit de parler de l'unicité. ■

Vu le théorème de Lagrange, on éprouve une tentation de chercher dans G un sous-groupe d'ordre m pour tout diviseur m de l'ordre n du groupe G. Or, en général, il n'y a pas de raisons de le faire. On peut vérifier que le groupe alterné A_4 d'ordre 12 ne contient pas de sous-groupes d'ordre 6.

Mais dans certains groupes « l'inversion du théorème de Lagrange » est vraie. Par exemple on a le théorème suivant:

THÉORÈME 6. — *Tout sous-groupe d'un groupe cyclique est encore un groupe cyclique. Les seuls sous-groupes du groupe cyclique infini* $(\mathbb{Z}, +)$ *gont les groupes (infinis)* $(m\mathbb{Z}, +)$, $m \in \mathbb{N}$. *Les sous-groupes d'un sroupe cyclique d'ordre* q *sont en correspondance biunivoque avec les diviseurs* d *(positifs) du nombre* q.

DÉMONSTRATION. — Considérons, pour varier le contenu, un groupe cyclique arbitraire $A = \langle a \rangle$ noté additivement. Chacun de ses éléments a donc la forme ka, où $k \in \mathbb{Z}$ ou bien $k = 0, 1, \ldots, q - 1$, si A est un groupe fini d'ordre q (voir théorème 3 du § 2). Soit B un sous-groupe non nul de A. Si $ka \in B$ pour un $k \neq 0$ quelconque, on a aussi $-ka \in B$. De tous les éléments $ka \in B$ avec k positif choisissons l'élément ma dont m est le plus petit.

Tout $k > 0$ s'écrit sous la forme $k = lm + r$, $0 \leqslant r < m$. Nous voyons que $ka \in B$ entraîne $ra = ka - l(ma) \in B$, c'est-à-dire $r = 0$. Donc, $B = \langle ma \rangle$ est un groupe cyclique.

Tous les groupes cycliques infinis sont isomorphes (théorème 1). Prenons à titre d'exemple le groupe additif $(\mathbb{Z}, +)$. Il est engendré par 1 ou -1, de sorte que, d'après ce qui a été démontré, tout sous-groupe de $(\mathbb{Z}, +)$ se définit par un entier naturel m et a la forme

$$m\mathbb{Z} = \langle m \cdot 1 \rangle = \{0, \pm m, \pm 2m, \ldots\}.$$

Tous ces sous-groupes sont évidemment infinis.

Soit maintenant $\langle a \rangle = \{0, a, \ldots, (q-1)a\}$, $qa = 0$. Nous savons que $B = \{0, ma, 2ma, \ldots\}$, où $m \in \mathbb{N}$, et $sa \in B$, $s \in \mathbb{N} \Rightarrow s = mt$. On affirme que m divise q. En effet, soit $q = dm + r$, $0 \leqslant r < m$. Alors

$$0 = qa = d(ma) + ra,$$

d'où $ra = -d(ma) \in B$. Le fait que m est minimal entraîne $r = 0$ et on a $q = dm$. Ainsi,

$$B = \{0, ma, 2ma, \ldots, (d-1)ma\} = mA$$

est un sous-groupe d'ordre d du groupe A. Lorsque m parcourt tous les diviseurs positifs du nombre q, il en est de même pour d et nous obtenons exactement un sous-groupe d'ordre d pour chaque diviseur d de q. ■

COROLLAIRE. — *Dans un groupe cyclique $\langle a \rangle$ d'ordre q le sous-groupe d'ordre $d \mid q$ coïncide avec l'ensemble des éléments $b \in \langle a \rangle$ tels que $db = 0$.*

DÉMONSTRATION. — Si $dm = q$ alors $b \in B = mA$ et $db = 0$. Réciproquement, soient $b = la \in \langle a \rangle$ et $db = 0$. Par hypothèse, $dla = 0$. Il en résulte que $dl = qk = dmk$, d'où $l = mk$ et $b = la = k(ma) \in mA$. ■

5. Monomorphisme $S_n \to \mathrm{GL}(n)$.— Rappelons qu'un monomorphisme des groupes $G \to G'$ est par définition un plongement isomorphe de G dans G'.

THÉORÈME 7. — *Il existe un monomorphisme $f: S_n \to \mathrm{GL}(n)$ tel que la matrice $f(\pi)$, $\pi \in S_n$ a pour déterminant $|f(\pi)| = \varepsilon_\pi$.*

DÉMONSTRATION. — Toute matrice (a_{ij}) de type (n, n) peut être présentée sous la forme d'une réunion des colonnes: $(a_{ij}) = (A^{(1)}, A^{(2)}, \ldots, A^{(n)})$. En particulier, soient

$$E^{(1)} = \begin{Vmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{Vmatrix}, \quad E^{(2)} = \begin{Vmatrix} 0 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{Vmatrix}, \quad \ldots, \quad E^{(n)} = \begin{Vmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{Vmatrix}$$

les colonnes de la matrice unité E. Définissons une application $f: S_n \to \mathrm{GL}(n)$ en posant

$$\pi \mapsto f(\pi) = (E^{(\pi(1))}, E^{(\pi(2))}, \ldots, E^{(\pi(n))}). \tag{1}$$

Ainsi, $f(\pi)$ est une matrice $n \times n$ dont chaque ligne et chaque colonne comporte exactement une unité, les autres éléments étant nuls. On conçoit sans peine que $f(\pi) \in \mathrm{GL}(n)$.

Soient σ, τ des permutations arbitraires et $\pi = \sigma\tau$ leur produit. Par définition, dans la i-ième ligne de la matrice $f(\sigma) = (a_{is})$ et dans la j-ième colonne de la matrice $f(\tau) = (b_{kj})$, les éléments non nuls sont égaux respectivement à $a_{i,\sigma^{-1}(i)} = 1$ et $b_{\tau(j), j} = 1$. C'est pourquoi, pour la matrice $f(\sigma) f(\tau) = (c_{ij})$, la condition $c_{i,j} \neq 0$ est équivalente à $\sigma^{-1}(i) = \tau(j)$, c'est-à-dire, $i = \sigma\tau(j) = \pi(j)$. Or, cela signifie justement que $f(\sigma) f(\tau) = f(\sigma\tau)$, donc f est un homomorphisme.

La propriété $\operatorname{Ker} f = e$ est évidente. En effet, par suite de (1), il est immédiat que $f(\pi) = E \Rightarrow \pi = e$. Donc, f est un monomorphisme.

Enfin, nous savons que le déterminant est une fonction symétrique gauche de ses colonnes. Par conséquent, $|f(\pi)| = g(E^{(1)}, \ldots, E^{(n)})$ est une fonction symétrique gauche des arguments $E^{(1)}, \ldots, E^{(n)}$. La relation (1), ainsi que la définition (5) du § 2 et la démonstration du théorème 5 du § 2 entraînent :

$$\varepsilon_\pi |f(\pi)| = \varepsilon_\pi \cdot g(E^{(1)}, \ldots, E^{(n)}) = (\pi \circ g)(E^{(1)}, \ldots, E^{(n)}) =$$
$$= g(E^{(\pi^{-1}(1))}, \ldots, E^{(\pi^{-1}(n))}) = |E^{(1)}, \ldots, E^{(n)}| = \det E = 1.$$

Donc, $|f(\pi)| = \varepsilon_\pi$. ■

Les matrices de la forme $f(\pi)$, $\pi \in S_n$, sont appelées *matrices des permutations*. La restriction à A_n du monomorphisme f est un monomorphisme dans SL $(n, \mathbb{R})$. La composition $f \circ L$ des applications $L: G \to S_n$ (théorème 2) et $f: S_n \to \mathrm{GL}(n)$ conduit au monomorphisme $G \to \mathrm{GL}(n)$ pour tout groupe fini G. Dans le cas de S_3, l'application f se présente sous la forme :

$$e \mapsto \begin{Vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{Vmatrix}, \quad (12) \mapsto \begin{Vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{Vmatrix}, \quad (13) \mapsto \begin{Vmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{Vmatrix},$$
$$(23) \mapsto \begin{Vmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{Vmatrix}, \quad (123) \mapsto \begin{Vmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{Vmatrix}, \quad (132) \mapsto \begin{Vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{Vmatrix}.$$

En utilisant le théorème 7, on démontre aisément le *théorème de développement complet d'un déterminant.*

THÉORÈME 8. — *Le déterminant*

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}$$

peut s'écrire sous la forme d'une somme algébrique de n! produits (appelés termes du déterminant) :

$$\det A = \sum_{\pi \in S_n} \varepsilon_\pi a_{\pi(1),1} a_{\pi(2),2} \ldots a_{\pi(n),n}. \tag{2}$$

DÉMONSTRATION. — Désignons par $|A^{(1)}, \ldots, A^{(j-1)}, E^{(i)}, A^{(j+1)}, \ldots, A^{(n)}|$ le déterminant obtenu à partir de $|A|$ en remplaçant la colonne $A^{(j)}$, indice j, par la colonne $E^{(i)}$ de la matrice unité. La formule donnant le développement du déterminant sui-

vant la j-ième colonne montre que pour le cofacteur A_{ij} de l'élément a_{ij} du déterminant $|A| = |A^{(1)}, \ldots, A^{(n)}|$ il existe une expression sous la forme d'un déterminant d'ordre n:

$$A_{ij} = |A^{(1)}, \ldots, A^{(j-1)}, E^{(i)}, A^{(j+1)}, \ldots, A^{(n)}|,$$

d'où, d'après la même formule, on obtient

$$\det A = \sum_i a_{ij}A_{ij} = \sum_i a_{ij} |A^{(1)}, \ldots, A^{(j-1)}, E^{(i)}, A^{(j+1)}, \ldots, A^{(n)}|.$$

Si l'on applique ce procédé tout d'abord pour $j = 1$, ensuite (à chacun des n termes) pour $j = 2$, etc., on obtient pour $\det A$ des expressions contenant respectivement n, n^2 et, enfin, n^n déterminants:

$$\det A = \sum_{i_1} a_{i_1,1} |E^{(i_1)}, A^{(2)}, \ldots, A^{(n)}| =$$
$$= \sum_{i_1} a_{i_1,1} \sum_{i_2} a_{i_2,2} |E^{(i_1)}, E^{(i_2)}, A^{(3)}, \ldots, A^{(n)}| =$$
$$= \sum_{i_1, i_2} a_{i_1,1}a_{i_2,2} |E^{(i_1)}, E^{(i_2)}, \ldots, A^{(n)}| = \ldots$$
$$\ldots = \sum_{i_1, i_2, \ldots, i_n} a_{i_1,1}a_{i_2,2} \ldots a_{i_n,n} |E^{(i_1)}, E^{(i_2)}, \ldots, E^{(i_n)}|.$$

Ici, $i_1, i_2, \ldots, i_n$ parcourent toutes les collections possibles des nombres $1, 2, \ldots, n$ (y compris celles, où il y a des répétitions). Utilisons toutes les n^n applications distinctes $\pi: \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ (voir § 1, nº 2, exemple 1)), où $\pi(1) = i_1, \ldots, \pi(n) = i_n$, et mettons l'expression de $\det A$ sous la forme

$$\det A = \sum_{\pi} a_{\pi(1),1}\, a_{\pi(2),2} \ldots a_{\pi(n),n} |E^{(\pi(1))}, E^{(\pi(2))}, \ldots, E^{(\pi(n))}|$$

(la somme est étendue à tous les π). Il reste à remarquer que si $\pi(i) = \pi(j)$ pour deux indices distincts quelconques i et j, les deux colonnes du déterminant $|E^{(\pi(1))}, \ldots, E^{(\pi(n))}|$ coïncident et donc le déterminant est nul. Par conséquent, le déterminant $|E^{(\pi(1))}, \ldots, E^{(\pi(n))}|$ est non nul si, et seulement si, l'application π est bijective, c'est-à-dire si elle est une permutation. Or, dans un tel cas, on a d'après le théorème 7 $|E^{(\pi(1))}, \ldots, E^{(\pi(n))}| = |f(\pi)| = \varepsilon_\pi$. ■

REMARQUE. — Il n'est pas certes difficile de démontrer le théorème 8 directement par récurrence sur n, sans aucune mention des groupes, bien que la nature des signes de ε_π devant les termes du déterminant relève finalement de la théorie des groupes. Le théorème 7 présente un intérêt indépendant.

On attire l'attention sur le fait que le théorème 8 peut servir de base pour la construction d'une théorie des déterminants (d'ailleurs souvent on l'utilise comme base). A savoir, en définissant le déterminant $\det A$ par la formule (2), nous aurions pu obtenir toutes ses propriétés, y compris la formule du développement de $\det A$ suivant les éléments de la première (ou j-ième) colonne, qui nous a servi de point de départ au chapitre 3.

EXERCICES

1. Démontrer qu'il n'existe, à un isomorphisme près, qu'un nombre fini $\rho(n)$ de groupes d'ordre donné n. (Indication. Trouver le majorant du nombre de différentes tables de Cayley d'ordre n. Le raisonnement formel utilisant le théorème 2 limite $\rho(n)$ par le nombre $\binom{n!}{n}$ de divers sous-ensembles de S_n à n éléments. En réalité, $\rho(n)$ est sensiblement plus petit, mais on n'a pas encore réussi à trouver une bonne estimation de la valeur réelle.)

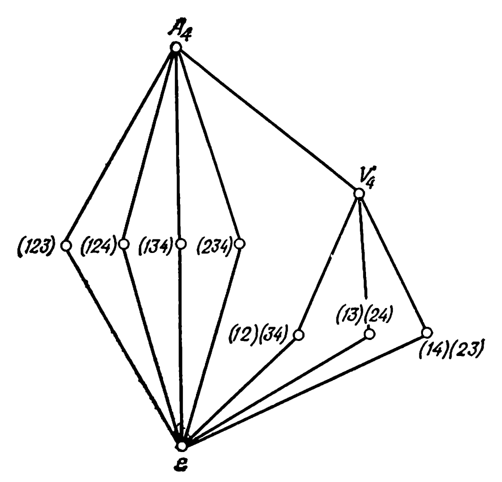


Fig. 14.

2. En utilisant l'exercice 7 du § 2, montrer que tout groupe fini peut être plongé dans un groupe fini engendré par deux éléments (autrement dit, montrer qu'il existe un monomorphisme).

3. Démontrer que dans tout groupe le sous-groupe d'indice 2 est nécessairement distingué. (Indication. Effectuer la partition de G d'abord en classes à gauche suivant H et ensuite en classes à droite ayant les mêmes représentants.)

4. En utilisant l'exercice 3, démontrer que S_3 est, à un isomorphisme près, l'unique groupe non commutatif d'ordre 6.

5. S'assurer que le diagramme (fig. 14) représente tous les sous-groupes du groupe alterné A_4. Le symbole V_4 désigne le groupe *à quatre éléments* (*groupe de Klein*) $V_4 = \{e, (12)(34), (13)(24), (14)(23)\}$; près des autres sommets du diagramme sont placés les éléments qui engendrent des sous-groupes cycliques.

6. Montrer que tous les groupes d'ordre 4 sont commutatifs. Les seuls groupes qui les représentent sont, à un isomorphisme près, les groupes des permutations $U = \langle(1234)\rangle$, V_4 ou ceux des matrices

$$L_1 = \left\{ \left\|\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}\right\|, \left\|\begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix}\right\|, \left\|\begin{matrix} -1 & 0 \\ 0 & -1 \end{matrix}\right\|, \left\|\begin{matrix} 0 & -1 \\ 1 & 0 \end{matrix}\right\| \right\} \subset \mathrm{GL}(2, \mathbb{R}),$$

$$L_2 = \left\{ \left\|\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}\right\|, \left\|\begin{matrix} 1 & 0 \\ 0 & -1 \end{matrix}\right\|, \left\|\begin{matrix} -1 & 0 \\ 0 & 1 \end{matrix}\right\|, \left\|\begin{matrix} -1 & 0 \\ 0 & -1 \end{matrix}\right\| \right\} \subset \mathrm{GL}(2, \mathbb{R}).$$

Ecrire sous une forme explicite les isomorphismes $U \to L_1$, $V_4 \to L_2$. (Indication. Si $x^2 = e$ pour tout élément $x \in G$, alors

$$abab = e \Rightarrow ab = b^{-1}a^{-1} = b\,(b^{-1})^2\,(a^{-1})^2 a = beea = ba.)$$

§ 4. Anneaux et corps

1. Définition et propriétés générales des anneaux.— Les structures algébriques $(\mathbb{Z}, +)$, $(\mathbb{Z}, \cdot)$ ont joué pour nous le rôle de tout premiers exemples de monoïdes. Quant à $(\mathbb{Z}, +)$ nous l'avons considéré plus tard comme un groupe additif commutatif (au fait, un groupe cyclique). Pourtant, dans la vie courante, ces structures sont le plus souvent réunies ensemble pour donner un être qu'on appelle en mathématiques anneau. Une des lois importantes de l'arithmétique élémentaire est la distributivité: $(a + b)\,c = ac + bc$ qui ne semble triviale qu'en raison de l'habitude acquise. Si l'on cherche à réunir, par exemple, les structures algébriques $(\mathbb{Z}, +)$, $(\mathbb{Z}, \circ)$, où $n \circ m = n + m + nm$, on ne constatera plus un si bon accord entre les deux opérations binaires. Avant de passer à d'autres exemples, donnons la définition de l'anneau.

DÉFINITION. — *Soit K un ensemble non vide muni de deux opérations* (*algébriques binaires*) *notées* $+$ (*addition*) et $\cdot$ (*multiplication*), *qui satisfont aux conditions suivantes*:

(K1) $(K, +)$ *est un groupe abélien*;

(K2) $(K, \cdot)$ *est un semi-groupe*;

(K3) *les opérations d'addition et de multiplication sont liées entre elles par les lois distributives* (*en d'autres termes*: *la multiplication est distributive par rapport à l'addition*):

$$(a + b)\,c = ac + bc, \qquad c\,(a + b) = ca + cb$$

quels que soient $a, b, c \in K$.

Alors, $(K, +, \cdot)$ *est appelé anneau.*

La structure $(K, +)$ s'appelle *groupe additif de l'anneau*, et $(K, \cdot)$ s'appelle *demi-groupe multiplicatif* de l'anneau. Si $(K, \cdot)$ est un monoïde, on dit que $(K, +, \cdot)$ est un *anneau unitaire*.

On convient de désigner l'élément unité de l'anneau par l'unité ordinaire 1. L'existence de 1 est souvent introduite dans la définition de l'anneau, mais nous ne le ferons pas.

Dans les applications et dans la théorie générale des anneaux (une telle théorie existe et à l'état bien développé) on considère des systèmes algébriques dans lesquels l'axiome (K2) est, soit complètement éliminé, soit remplacé par un autre axiome suivant le problème concret. Dans de tels cas on parle des *anneaux non associatifs*. Dans le présent chapitre nous ne nous proposons d'étudier que des anneaux ordinaires (*associatifs*). Cela signifie que nous pourrons nous appuyer sur le théorème 1 du § 1 sans nous soucier des parenthèses dans le produit $a_1a_2 \ldots a_k$, quel que soit le nombre k d'éléments de l'anneau.

Toute partie L de l'anneau K s'appelle *sous-anneau* si

$$x, y \in L \Rightarrow x - y \in L \quad \text{et} \quad xy \in L,$$

c'est-à-dire, si L est un sous-groupe du groupe additif et un sous-demi-groupe du demi-groupe multiplicatif de l'anneau.

Il est clair que l'intersection de toute famille de sous-anneaux de K est un sous-anneau (les raisonnements sont les mêmes que dans le cas des groupes) et donc il y a tout lieu de parler d'un sous-anneau $\langle T \rangle \subset K$ *engendré par un sous-ensemble* $T \subset K$. Par définition $\langle T \rangle$ est l'intersection de tous ceux des sous-anneaux de K qui contiennent T. Si T était dès le début un sous-anneau, alors $\langle T \rangle = T$.

Un anneau est dit *commutatif*, si $xy = yx$, quels que soient x, $y \in K$ (à la différence des groupes, dans les ouvrages mathématiques russes l'anneau commutatif ne s'appelle pas anneau abélien!)

La notion d'anneau, telle que nous l'avons introduite, est une notion très large. De plus, la classe d'anneaux commutatifs, qui semble au premier abord assez spéciale, a été, pendant plusieurs décennies, l'objet d'études intenses, et à présent la théorie des anneaux commutatifs s'enchevêtre avec la géométrie algébrique, une discipline mathématique élégante qui est à la charnière de l'algèbre, de la géométrie et de la topologie.

EXEMPLES. 1) Soit $(\mathbb{Z}, +, \cdot)$ l'*anneau des entiers relatifs*, où sont définies les opérations ordinaires d'addition et de multiplication. L'ensemble $m\mathbb{Z}$ des entiers relatifs, divisibles par m, est un sous-anneau de $\mathbb{Z}$ (sans élément unité pour $m > 1$). De même, $\mathbb{Q}$ et $\mathbb{R}$ sont les anneaux unitaires; ceci étant les inclusions $\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$ définissent une suite de sous-anneaux de l'anneau $\mathbb{R}$.

2) Les propriétés des opérations d'addition et de multiplication dans $M_n(\mathbb{R})$ que nous avons introduites et étudiées en détail au chapitre 2, permettent d'affirmer que $M_n(\mathbb{R})$ est un anneau possédant un élément unité $1 = E$. Il s'appelle *anneau complet des matrices sur* $\mathbb{R}$ ou, encore, *anneau des matrices carrées d'ordre n* (ou *de type* (n, n) *sur* $\mathbb{R}$. C'est un des plus importants exemples d'anneaux. Etant donné que pour $n > 1$ les matrices ne sont pas en général permutables, $M_n(\mathbb{R})$ est un anneau non commutatif. Il contient comme sous-anneaux les anneaux $M_n(\mathbb{Q})$ et $M_n(\mathbb{Z})$ des matrices carrées du même ordre respectivement sur $\mathbb{Q}$ et sur $\mathbb{Z}$. En général, $M_n(\mathbb{R})$ contient un très grand nombre de divers sous-anneaux. Nous verrons que de temps en temps certains d'entre eux apparaîtront de façon naturelle. Remarquons encore qu'on peut considérer un anneau des matrices carrées $M_n(K)$ sur un anneau commutatif arbitraire K, car l'addition et la multiplication de deux matrices A, $B \in M_n(K)$ donnent une nouvelle matrice à coefficients dans K, et les lois de distributivité dans $M_n(K)$ sont conséquences des lois analogues dans K. Tout cela découle directement des opérations formelles sur les matrices que nous avons résumées à la fin des n°s 1 et 3 du chapitre 2, § 3.

3) L'*anneau des fonctions* est utilisé dans les différentes branches des mathématiques aussi largement que l'anneau des matrices. A savoir, soient X un ensemble et K un anneau quelconques. Considérons l'ensemble $K^X = \{X \to K\}$ de toutes les fonctions (applications) $f: X \to K$ muni de deux opérations binaires appelées *addition* et *multiplication* des fonctions et définies comme suit:

$$(f + g)(x) = f(x) \oplus g(x),$$
$$(fg)(x) = f(x) \odot g(x)$$

($\oplus$ et $\odot$ sont les opérations d'addition et de multiplication dans K). Il est évident que nous n'avons pas ici la même composition des fonctions qui nous

a conduit dans le cas des applications linéaires à l'anneau M_n. Nous choisissons plutôt un point de vue adopté en Analyse mathématique. Etant donné, par exemple $X = \mathbb{R}$ et $K = \mathbb{R}$, le produit des fonctions tg et sin sera tg·sin: $x \mapsto \mathrm{tg}\, x \cdot \sin x$ et non tg ∘ sin: $x \mapsto \mathrm{tg}\,(\sin x)$.

On vérifie sans peine que K^X satisfait à tous les axiomes d'anneau. Par exemple, vu la distributivité des opérations dans K, on a

$$[f(x) \oplus g(x)] \odot h(x) = f(x) \odot h(x) \oplus g(x) \odot h(x)$$

quelles que soient les fonctions f, g, $h \in K^X$ et $x \in X$, ce qui donne, par la définition même des opérations introduites, $(f + g)\, h = fh + gh$. On établit de même la vérité de la deuxième loi distributive. Si 0, 1 sont les éléments zéro et unité de K, alors

$$0_X: x \mapsto 0, \qquad 1_X: x \mapsto 1$$

sont des fonctions *constantes* qui jouent les rôles de zéro et d'unité dans K^X. Si K est commutatif, il en est de même de l'anneau des fonctions K^X.

L'anneau K^X contient de différents sous-anneaux définis par les propriétés spéciales des fonctions. Soit, par exemple $X = [0, 1]$ un intervalle fermé de $\mathbb{R}$ et $K = \mathbb{R}$. Alors, l'anneau $\mathbb{R}^{[0,\,1]}$ de toutes les fonctions réelles définies sur $[0, 1]$ contient parmi les sous-anneaux: l'anneau $\mathbb{R}^{[0,\,1]}_{\mathrm{bor}}$ de toutes les fonctions bornées, l'anneau $\mathbb{R}^{[0,\,1]}_{\mathrm{cont}}$ de toutes les fonctions continues, l'anneau $\mathbb{R}^{[0,\,1]}_{\mathrm{dér}}$ de toutes les fonctions continûment dérivables, etc., car toutes les propriétés indiquées se conservent lors de l'addition (de la soustraction) et de la multiplication des fonctions.

A chaque nombre $a \in \mathbb{R}$ correspond une fonction *constante* $a_X: x \mapsto a$, de sorte que l'application injective $a \mapsto a_X$ permet de considérer $\mathbb{R}$ comme sous-anneau de $\mathbb{R}^X$. Bref, presque toute classe naturelle de fonctions se trouve en correspondance avec un sous-anneau de $\mathbb{R}^X$.

4) La relation $xy = 0$ définie pour tout couple $x, y \in A$ confère à tout groupe additif abélien $(A, +)$ la structure d'*anneau à multiplication nulle.*

De nombreuses propriétés des anneaux ne sont que des énoncés modifiés des propriétés correspondantes des groupes et, en général, des ensembles munis d'une seule opération associative. Par exemple, on a $a^m a^n = a^{m+n}$, $(a^m)^n = a^{mn}$ pour tous les entiers non négatifs m, n et tout $a \in K$ (comparer avec la relation (2) du § 1). D'autres propriétés plus caractéristiques des anneaux, qui sont des conséquences directes des axiomes, simulent au fond les propriétés de $\mathbb{Z}$. Signalons certaines de ces dernières propriétés. Premièrement,

$$a \cdot 0 = 0 \cdot a = 0 \text{ pour tout } a \in K. \tag{1}$$

En effet, $a + 0 = a \Rightarrow a(a + 0) = aa \Rightarrow a^2 + a \cdot 0 = a^2 \Rightarrow a^2 + a \cdot 0 = a^2 + 0 \Rightarrow a \cdot 0 = 0$ (d'une manière analogue $0 \cdot a = 0$).

En supposant pour un instant que $0 = 1$, on obtient $a = a \cdot 1 = a \cdot 0 = 0$ pour tout $a \in K$, c'est-à-dire que K ne se compose que de zéro. Par conséquent, pour un anneau non trivial K, on a toujours $0 \neq 1$. On a aussi

$$(-a) \cdot b = a(-b) = -(ab), \tag{2}$$

car, par exemple, de (1) et de l'axiome de distributivité il résulte que

$$0 = a \cdot 0 = a(b - b) = ab + a(-b) \Rightarrow a(-b) = -(ab). \quad (3)$$

Comme $-(-a) = a$, on obtient de (2) les égalités $(-a)(-b) = ab$ (par exemple, $(-1)(-1) = 1$) et $-a = (-1) \cdot a$.

L'axiome de distributivité a pour conséquence la *loi de distributivité générale* :

$$(a_1 + \ldots + a_n)(b_1 + \ldots + b_m) = \sum_{i=1}^{n} \sum_{j=1}^{m} a_i b_j. \quad (4)$$

On le vérifie sans peine en raisonnant par récurrence d'abord sur n (pour $m = 1$) et ensuite sur m. En utilisant maintenant les relations (1), (2) et (3), on obtient

$$n(ab) = (na)b = a(nb)$$

pour tout $n \in \mathbb{Z}$ et tous les $a, b \in K$

Signalons enfin la formule du binôme de Newton

$$(a + b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}, \quad (5)$$

valable pour tous les $a, b \in K$, K étant un anneau commutatif. Pour démontrer (5) il faut, en s'appuyant sur (4), opérer de la même façon qu'au chapitre 1, § 7, où nous avons considéré le cas particulier, où $K = \mathbb{Z}$.

2. Congruences. Anneau des classes résiduelles.— D'après le théorème 6 du § 3, les sous-groupes non nuls du groupe $(\mathbb{Z}, +)$ sont représentés par les groupes $m\mathbb{Z}$, et eux seuls, où m parcourt l'ensemble $\mathbb{N}$ des entiers naturels. L'ensemble $m\mathbb{Z}$ est évidemment stable non seulement pour l'opération d'addition mais aussi pour la multiplication, en satisfaisant à tous les trois axiomes d'anneau. Ainsi, on a la proposition suivante : tout sous-anneau non nul de l'anneau $\mathbb{Z}$ est de la forme $m\mathbb{Z}$, où $m \in \mathbb{N}$.

Cherchons maintenant à construire, en utilisant le sous-anneau $m\mathbb{Z} \subset \mathbb{Z}$, un anneau non nul constitué d'un nombre fini d'éléments. A cet effet, introduisons la définition suivante :

DÉFINITION. — *On dit que deux entiers n, n' sont congrus modulo m si leur division par m donne mêmes restes. Dans ce cas on écrit $n \equiv n'$ (m) ou $n \equiv n'$* (mod. m) *et on appelle m module de congruence.*

On obtient ainsi une partition de $\mathbb{Z}$ en classes de nombres congrus modulo m et appelées *classes résiduelles* modulo m. Toute classe ré-

siduelle est de la forme

$$\{r\}_m = r + m\mathbb{Z} = \{r + mk \mid k \in \mathbb{Z}\},$$

si bien que

$$\mathbb{Z} = \{0\}_m \cup \{1\}_m \cup \ldots \cup \{m-1\}_m. \tag{6}$$

Nous remarquons que les classes résiduelles sont des classes du groupe additif $\mathbb{Z}$ suivant le sous-groupe $m\mathbb{Z}$ alors que la partition (6) correspond à la décomposition dont il s'agit dans le théorème 4 du § 2. Par définition, $n \equiv n'\ (m) \Leftrightarrow n - n'$ est divisible par m. La commodité de la notation $n \equiv n'\ (m)$ pour la relation de divisibilité $m \mid (n - n')$ réside en ce qu'avec de telles congruences on peut opérer tout à fait de la même façon qu'avec les égalités ordinaires. A savoir, si $k \equiv k'\ (m)$ et $l \equiv l'\ (m)$, on a $k \pm l \equiv k' \pm$ $\pm\ l'\ (m)$ et $kl' \equiv k'l'\ (m)$.

En particulier, $k \equiv k'\ (m) \Rightarrow ks \equiv k's\ (m)$ pour tout $s \in \mathbb{Z}$.

Ainsi, à deux classes quelconques $\{k\}_m$ et $\{l\}_m$ on peut faire associer, indépendamment des représentants choisis k et l, des classes qui sont leur somme et leur produit, ce qui signifie que sur l'ensemble $\mathbb{Z}_m = \mathbb{Z}/m\mathbb{Z}$ des classes résiduelles modulo m sont univoquement induites les opérations $\oplus$ et $\odot$:

$$\begin{aligned} \{k\}_m \oplus \{l\}_m &= \{k + l\}_m, \\ \{k\}_m \odot \{l\}_m &= \{kl\}_m. \end{aligned} \tag{7}$$

Puisque la définition de ces opérations se ramène à des opérations correspondantes sur les nombres des classes résiduelles, à savoir sur les éléments de $\mathbb{Z}$, l'anneau $\{Z_m, \oplus, \odot\}$ sera, lui aussi, un anneau commutatif unitaire, avec $\{1\}_m = 1 + m\mathbb{Z}$. Il s'appelle *anneau des classes résiduelles modulo m*. Lorsqu'on acquiert une certaine habitude (et le module est fixe) on omet l'indice m et on écrit $\bar{k}$ au lieu de $\{k\}_m$, si bien que

$$\bar{k} \oplus \bar{l} = \overline{k + l},$$

$$\bar{k} \odot \bar{l} = \overline{kl}.$$

La plus haute maîtrise de Z_m, qui semble au premier abord être de caractère d'un sacrilège, mais qui présente des avantages techniques évidents, est qu'on abandonne les symboles spéciaux (barres et cercles) pour opérer avec un ensemble des représentants modulo m, le plus souvent avec $\{0, 1, 2, \ldots, m-1\}$ que l'on appelle *système réduit des entiers modulo m*. Avec cette convention, on a par exemple : $-k = m - k$, $2(m-1) = -2 = m - 2$.

Ainsi, les anneaux finis existent. Donnons trois exemples des plus simples anneaux, en indiquant pour chacun les tables d'addition

et de multiplication:

Z_2:

+	0	1
0	0	1
1	1	0

·	0	1
0	0	0
1	0	1

Z_3:

+	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

·	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

Z_4:

+	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

·	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

L'anneau Z_m des classes résiduelles attirait depuis longtemps l'attention des spécialistes en théorie des nombres et servait, en algèbre, de point de départ pour toutes sortes de généralisations.

3. Homomorphismes et idéaux des anneaux.— L'application f: $n \mapsto \{n\}_m$ possède en vertu des relations (7) les propriétés suivantes: $f(k+l) = f(k) \oplus f(l)$, $f(kl) = f(k) \odot f(l)$. Cela nous permet de parler de l'homomorphisme de $\mathbb{Z}$ dans Z_m conformément à la définition générale.

DÉFINITION. — *Soient donnés les anneaux* $(K, +, \cdot)$ *et* $(K', \oplus, \odot)$. *On dit que l'application* f: $K \to K'$ *est un homomorphisme si elle respecte toutes les opérations, c'est-à-dire, si*

$$f(a+b) = f(a) \oplus f(b),$$
$$f(ab) = f(a) \odot f(b).$$

Dans ce cas on a, certes, $f(0) = 0'$ et $f(na) = nf(a)$, $n \in \mathbb{Z}$.

On appelle *noyau* de l'homomorphisme f l'ensemble

$$\operatorname{Ker} f = \{a \in K \mid f(a) = 0'\}.$$

Il est clair que $\operatorname{Ker} f$ est un sous-anneau de K. Mais ce sous-anneau est loin d'être arbitraire. En effet, si $L = \operatorname{Ker} f \subset K$, on a $L \cdot x \subseteq \subseteq L$ (car $f(lx) = f(l) \odot f(x) = 0' \odot f(x) = 0'$ pour tout $l \in L$) et $x \cdot L \subseteq L$ pour tout $x \in K$. Donc $LK \subset L$ et $KL \subset L$. Le sous-anneau vérifiant ces propriétés s'appelle *idéal* (bilatère) de l'anneau K. Ainsi, les noyaux des homomorphismes sont toujours des idéaux.

De même que dans le cas des groupes (voir § 3, n° 3, terminologie), l'homomorphisme f: $K \mapsto K'$ s'appelle *monomorphisme* si $\operatorname{Ker} f = 0$, *épimorphisme* si l'image coïncide avec K':

$$\operatorname{Im} f = f(K) = \{a' \in K' \mid a' = f(a)\} = K',$$

et *isomorphisme* si l'application f est monomorphe et épimorphe. L'isomorphisme des anneaux est noté $K \cong K'$.

L'application $f: n \mapsto \{n\}_m$ que nous avons considérée plus haut, est évidemment un épimorphisme $\mathbb{Z} \to Z_m$ avec le noyau $\operatorname{Ker} f = m\mathbb{Z}$. Lors de la construction de Z_m on a justement utilisé de façon implicite le fait que $m\mathbb{Z}$ est un idéal de l'anneau $\mathbb{Z}$. Nous voyons que tout sous-anneau non nul de l'anneau $\mathbb{Z}$ est un idéal. Cette propriété est due au caractère spécifique de $\mathbb{Z}$; elle n'est plus vraie, par exemple dans l'anneau des matrices $M_2(\mathbb{Z})$: l'ensemble

$$\left\{\begin{pmatrix} \alpha & \beta \\ 0 & \delta \end{pmatrix} \middle| \alpha, \beta, \delta \in \mathbb{Z}\right\}$$

est un sous-anneau mais non un idéal de $M_2(\mathbb{Z})$.

L'exemple de $m\mathbb{Z}$ suggère un procédé de construction des idéaux (peut-être non pour tous) dans un anneau *commutatif* arbitraire K : si a est un élément quelconque de K, l'ensemble aK est toujours un idéal de K. En effet

$$ax + ay = a(x + y), \quad (ax)y = a(xy).$$

On dit que aK est *idéal principal* engendré par l'élément $a \in K$.

Les idéaux de tout anneau sont des sous-groupes du groupe additif de l'anneau, stables pour la multiplication à gauche et à droite par les éléments de l'anneau, si l'on ne prend que des anneaux unitaires, il s'avère judicieux d'introduire dans la définition de l'homomorphisme $f: K \to K'$ la condition $f(1) = 1'$. Dans le cas de l'épimorphisme cette condition est, certes, automatiquement réalisée.

Les anneaux isomorphes s'identifient d'après leurs propriétés algébriques, de sorte que seules les propriétés, qui se conservent lors des applications isomorphes, présentent un vrai intérêt mathématique. C'est justement cette circonstance qu'on avait en vue en considérant l'anneau Z_m soit comme ensemble des classes résiduelles modulo m, soit comme ensemble des représentants de ces classes arbitrairement choisis.

4. Notions de groupe quotient et d'anneau quotient.— Les sous-groupes distingués des groupes et les idéaux des anneaux ont une origine commune, ils sont noyaux des homomorphismes. Cette circonstance trouve son expression dans le fait général du passage au quotient, que nous allons analyser brièvement. Cette question sera étudiée de façon beaucoup plus détaillée dans la deuxième partie du livre.

Commençons par les groupes. La relation d'équivalence $\sim$ dans un groupe G définie par la partition de G en classes suivant un groupe distingué H possède une propriété remarquable. A savoir, si a et b sont deux éléments quelconques du groupe G et si $a \sim c$, $b \sim d$, alors par la définition (voir démonstration du théorème 4 du § 3) on a $a^{-1}c = h_1 \in H$, $b^{-1}d = h_2 \in H$, d'où

$$(ab)^{-1}cd = b^{-1}a^{-1}cd = b^{-1}(a^{-1}c)d = b^{-1}h_1b(b^{-1}d) = h_1'h_2 \in H$$

et donc $ab \sim cd$. On a utilisé ici la propriété que possède le sous-groupe distingué H de G: $b^{-1}h_1b = h_1' \in H$. Ainsi,

$$a \sim c, \quad b \sim d \Rightarrow ab \sim cd.$$

Cela signifie au fait que l'opération de multiplication définie dans le groupe G, induit une opération de multiplication sur l'ensemble quotient $G/\sim$ (voir chap. 1, § 6, nº 3) que nous sommes convenus de désigner par le symbole G/H.

Il y a intérêt à parler du produit des sous-ensembles arbitraires A, B du groupe G en entendant par AB l'ensemble de tous les produits ab, avec $a \in A$, $b \in B$. L'associativité dans G entraîne la relation

$$(AB)\,C = \{(ab)\,c\} = \{a\,(bc)\} = A\,(BC),$$

si bien qu'un sous-ensemble $H \subset G$ est sous-groupe de G si, et seulement si, $H^2 = H$, $H^{-1} = \{h^{-1} \mid h \in H\} \subset H$.

De ce point de vue, on interprète la classe aH comme produit de l'ensemble à un élément $\{a\}$ par le sous-groupe H. Le produit des classes aH, bH est un ensemble $aH \cdot bH$ qui, en général, n'est pas nécessairement une classe suivant H. Ainsi, la partition de S_3 suivant $H = \{e, (12)\}$ que nous avons considérée au § 3, nº 4, montre par exemple que

$$H \cdot (13)\,H = (13)\,H \cup (23)\,H.$$

Il n'en est plus de même lorsque H est un sous-groupe distingué du groupe G. Puisque $gH = Hg$ pour tout $g \in G$, on a

$$aH \cdot bH = a\,(Hb)\,H = a\,(bH)\,H = abH^2 = abH,$$

si bien que le raisonnement développé plus haut montre que la classe abH ne dépend pas des représentants a, b des classes aH, bH. Les propriétés

$$aH \cdot bH = abH,$$
$$H \cdot aH = aH \cdot H = aH,$$
$$a^{-1}H \cdot aH = aH \cdot a^{-1}H = eH = H$$

permettent d'énoncer le théorème suivant:

THÉORÈME 1. — *Si H est un sous-groupe distingué d'un groupe G, l'opération de multiplication $aH \cdot bH = abH$ confère à l'ensemble quotient G/H une structure du groupe appelé groupe quotient de G par H. La classe H est l'élément unité de G/H, et $a^{-1}H = (aH)^{-1}$ est l'élément inverse de aH.* ∎

Lorsque le groupe G est fini, l'ordre du groupe quotient G/H est défini par la formule

$$|G/H| = \frac{|G|}{|H|} = (G:H),$$

qui n'a rien d'inattendu après tout ce qui vient d'être dit, y compris le théorème de Lagrange (§ 3, nº 4).

Dans le cas des groupes abéliens notés additivement, l'opération binaire sur G/H est introduite par la relation

$$(a + H) + (b + H) = (a + b) + H,$$

et G/H est souvent appelé groupe G modulo H. Lorsque $G = \mathbb{Z}$, $H = m\mathbb{Z}$, on dit aussi « groupe $\mathbb{Z}$ modulo m ».

En passant à la construction de l'*anneau quotient* K/L de l'anneau K par l'idéal L nous adoptons que la « base » de l'anneau est constituée par un groupe additif abélien. C'est pourquoi, il convient de prendre pour éléments de K/L les classes $a + L$ (appelées *classes résiduelles modulo* L) dont l'addition se fait d'après la règle

$$(a + L) \oplus (b + L) = (a + b) + L, \tag{8}$$
$$\ominus (a + L) = -a + L,$$

et le produit se définit par

$$(a + L) \odot (b + L) = ab + L. \tag{9}$$

Il importe de s'assurer que la multiplication ainsi définie soit correcte, c'est-à-dire qu'elle ne dépende pas des représentants choisis dans les classes correspondantes. Soient $a' = a + x$, $b' = b + y$, où $x, y \in L$. Alors

$$a'b' = ab + ay + xb' = ab + z,$$

où $z = ay + xb' \in L$ car L est un idéal bilatère. Voilà pourquoi $a'b'$ se trouve dans la même classe d'équivalence que l'élément ab, ce qui signifie justement que le produit (9) est défini correctement. Pour abréger la notation, posons $\bar{a} = a + L$, de sorte que

$$\bar{a} \oplus \bar{b} = \overline{a + b}, \quad \bar{a} \odot \bar{b} = \overline{ab}.$$

En particulier, $\bar{0} = L$ et $\bar{1} = 1 + L$ (si K admet un élément unité 1). Il faut encore s'assurer que l'ensemble $\bar{K} = K/L = \{\bar{a} \mid a \in K\}$ muni des opérations $\oplus$, $\odot$ satisfait à tous les axiomes d'anneau. Or, cela est assez évident, car les opérations sur les classes résiduelles de $\bar{K}$ se ramènent aux opérations sur les éléments de K. Par exemple, la distributivité est vérifiée comme suit

$$(\bar{a} \oplus \bar{b}) \odot \bar{c} = (\overline{a + b}) \odot \bar{c} = \overline{(a + b)\, c} = \overline{ac + bc} =$$
$$= \overline{ac} \oplus \overline{bc} = \bar{a} \odot \bar{c} \oplus \bar{b} \odot \bar{c}.$$

Tout cela montre que l'application

$$\pi : a \mapsto \bar{a}$$

des anneaux $K \to K'$ est un épimorphisme avec le noyau $\operatorname{Ker} \pi = L$.

Ainsi, l'exemple de l'anneau quotient $Z_m = \mathbb{Z}/m\mathbb{Z}$ et de l'épimorphisme $\mathbb{Z} \to Z_m$ nous conduit à une situation analogue dans les anneaux quelconques.

Il convient de remarquer, bien que cela sorte du cadre de notre but immédiat (explication de la structure de Z_m au point de vue de l'algèbre générale) que toutes les images de l'anneau K par des homomorphismes sont en réalité les anneaux quotients de K et eux seuls, par des idéaux correspondants. En effet, si $f: K \to K'$ est un homomorphisme et $f(K)$ est l'image de K par f, nous obtenons un épimorphisme à condition de considérer $f(K) \subset K'$ au lieu de K'. Pour ne pas compliquer les notations, supposons dès le début que f est un épimorphisme, c'est-à-dire posons $f(K) = K'$. Conformément au principe général exposé au chapitre 1, § 6, nº 3, f définit une relation d'équivalence O_f sur K; dans le cas considéré, O_f est définie par la partition de K en classes $a + \operatorname{Ker} f = C_a$. L'application f engendre une bijection f' entre les éléments $a' \in K'$ et les classes C_a, à savoir $f'(C_a) = a'$ si $a' = f(a)$. Ceci étant,

$$f'(C_a + C_b) = f'(C_{a+b}) = f(a+b) = f(a) + f(b) =$$
$$= f'(C_a) + f'(C_b),$$
$$f'(C_a \cdot C_b) = f'(C_{ab}) = f(ab) = f(a) \cdot f(b) = f'(C_a) \cdot f'(C_b),$$

si bien que l'application bijective f' est un isomorphisme (pour simplifier, les opérations d'addition et de multiplication dans K, dans l'anneau des classes résiduelles $K/\operatorname{Ker} f$ et dans K' sont notées identiquement: $+$ et $\cdot$). Ainsi nous avons démontré le théorème suivant:

THÉORÈME 2 (théorème fondamental d'homomorphie des anneaux). — *Tout idéal L d'un anneau K définit* (à l'aide des formules (8), (9)) *sur l'ensemble quotient K/L une structure d'anneau, K/L étant l'image de l'anneau K par l'homomorphisme de noyau L. Réciproquement, toute image homomorphe $K' = f(K)$ d'un anneau K est isomorphe à l'anneau quotient $K/\operatorname{Ker} f$.*

REMARQUE.— Le second membre de la formule (9) ne coïncide pas en général avec le produit, dans le sens ensembliste, des classes résiduelles $a + L$ et $b + L$. Par exemple, pour $K = \mathbb{Z}$, $L = 8\mathbb{Z}$, le nombre entier $24 \in 16 + 8\mathbb{Z}$ n'est pas contenu dans $(4 + 8\mathbb{Z})^2$ car $(4 + 8s)(4 + 8t) = 16u$.

5. Types d'anneaux. Corps.— Dans les anneaux numériques bien connus $\mathbb{Z}$, $\mathbb{Q}$ et $\mathbb{R}$, $ab = 0$ implique soit $a = 0$, soit $b = 0$. Cette propriété n'est pourtant plus vraie dans l'anneau M_n des matrices carrées. En utilisant les matrices E_{ij} (voir chap. 2, § 3, démonstration du théorème 3) nous obtenons les égalités $E_{ij}E_{kl} = 0$ pour $j \neq k$, bien que, certes, $E_{ij} \neq 0$ et $E_{kl} \neq 0$. Remarquons que $E_{ik}E_{kj} = E_{ij} \neq 0$. On pourrait attribuer ce phénomène, si étrange pour l'arithmétique élémentaire, à la non-commutativité de l'anneau M_n, mais il n'en est pas ainsi. Comme nous l'avons vu au nº 2, les élé-

ments de l'anneau commutatif $\mathbb{Z}_4$ vérifient l'égalité $2 \odot 2 = 0$ en dépit de la vérité connue: « deux fois deux font quatre ».

Voilà encore deux exemples.

EXEMPLE 1. — L'ensemble des couples numériques (a, b) $(a, b \in \mathbb{Z}$, $\mathbb{Q}$ ou $\mathbb{R})$ muni des opérations d'addition et de multiplication définies par les formules

$$(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2),$$
$$(a_1, b_1)\cdot(a_2, b_2) = (a_1a_2, b_1b_2),$$

est manifestement un anneau commutatif avec l'élément unité $(1, 1)$, dans lequel nous observons de nouveau le même phénomène: $(1, 0)\cdot(0, 1) = (0, 0) = 0$.

EXEMPLE 2. — Dans l'anneau $\mathbb{R}^{\mathbb{R}}$ des fonctions réelles (voir nº 1, exemple 3) les fonctions $f: x \mapsto |x| + x$ et $g: x \mapsto |x| - x$ sont telles que $f(x) = 0$ pour $x \leqslant 0$ et $g(x) = 0$ pour $x \geqslant 0$. Leur produit fg est de ce fait une fonction nulle, bien que $f \neq 0$ et $g \neq 0$.

DÉFINITION. — *Si, dans un anneau K, $ab = 0$ avec $a \neq 0$ et $b \neq 0$, on dit que a est un diviseur de zéro à gauche et b un diviseur de zéro à droite (si l'anneau K est commutatif, on parle tout simplement des diviseurs de zéro). Le zéro lui-même est, dans un anneau $K \neq 0$, un diviseur trivial de zéro. S'il n'y a pas d'autres diviseurs de zéro (sauf 0) on dit que K est un anneau sans diviseurs de zéro. Un anneau commutatif sans diviseurs de zéro et ayant un élément unité $1 \neq 0$ est appelé anneau intègre (on dit quelquefois anneau d'intégrité ou domaine d'intégrité).*

THÉORÈME 3. — *Un anneau commutatif unitaire non trivial K est intègre si, et seulement si, la loi de simplification*

$$ab = ac, \quad a \neq 0 \Rightarrow b = c$$

est vérifiée quels que soient $a, b, c \in K$.

En effet, si K est muni de la loi de simplification, $ab = 0 = a\cdot 0$ entraîne soit $a = 0$, soit $a \neq 0$ avec $b = 0$. Inversement, si K est intègre, on a $ab = ac$, $a \neq 0 \Rightarrow a(b - c) = 0$, $b - c = 0 \Rightarrow b = c$. ■

Il est naturel de considérer dans un anneau unitaire K l'ensemble des éléments inversibles: un élément a s'appelle *inversible* (ou *diviseur d'unité*) s'il existe un élément a^{-1} tel que $aa^{-1} = 1 = a^{-1}a$. D'une manière plus précise, il faudrait parler des éléments *inversibles à droite* ou *à gauche* ($ab = 1$ ou $ba = 1$), mais dans les anneaux commutatifs, de même que dans les anneaux sans diviseurs de zéro, ces notions coïncident. En effet, il résulte de $ab = 1$ que $aba = a$, d'où $a(ba - 1) = 0$. Puisque $a \neq 0$, on a $ba - 1 = 0$, c'est-à-dire $ba = 1$.

Nous savons que, par exemple, *dans l'anneau M_n les éléments inversibles sont les matrices de déterminant non nul et elles seules.* Un élément inversible a ne peut pas être un diviseur de zéro: $ab = 0 \Rightarrow a^{-1}(ab) = 0 \Rightarrow (a^{-1}a)b = 0 \Rightarrow 1\cdot b = 0 \Rightarrow b = 0$ (d'une manière

analogue, $ba = 0 \Rightarrow b = 0$). Il ne faut donc pas s'étonner de voir l'assertion suivante :

THÉORÈME 4. — *Tous les éléments inversibles d'un anneau unitaire K forment un groupe $U(K)$ pour la multiplication.*

En effet, puisque l'ensemble $U(K)$ possède un élément unité, $a \in U(K) \Rightarrow a^{-1} \in U(K)$ et la multiplication est associative dans K, il ne reste qu'à nous assurer que l'ensemble $U(K)$ est stable pour la multiplication, c'est-à-dire vérifier que le produit ab de deux éléments quelconques a et b de $U(K)$ appartient encore à $U(K)$. Or, cela est évident car $(ab)^{-1} = b^{-1}a^{-1}$ $(ab \cdot b^{-1}a^{-1} = a(bb^{-1})a^{-1} = a \cdot 1 \cdot a^{-1} = aa^{-1} = 1)$, et donc ab est inversible. ■

On établit sans peine que $U(\mathbb{Z}) = \{\pm 1\}$ est un groupe cyclique d'ordre 2.

Si, dans la définition de l'anneau, on remplace l'axiome (K2) par une condition sensiblement plus forte (K2′) : l'ensemble $K^* = K \setminus \{0\}$ est un groupe pour la multiplication, on obtient une classe d'anneaux très intéressante appelés *corps*. Le corps est donc un anneau sans diviseurs de zéro, dont tout élément non nul est inversible. Lorsque la multiplication est commutative, le corps est dit *commutatif*, et les opérations d'addition et de multiplication deviennent presque complètement symétriques. Ainsi, donnons encore une fois la définition du corps commutatif.

DÉFINITION. — *On appelle corps commutatif P un anneau commutatif ayant un élément unité $1 \neq 0$ et dont tout élément $a \neq 0$ est inversible. Le groupe $P^* = U(P)$ s'appelle groupe multiplicatif de P.*

Le corps commutatif peut être considéré comme une combinaison de deux groupes abéliens, additif et multiplicatif, liés par une loi de distributivité (maintenant une seule, étant donné la commutativité). Le produit ab^{-1} s'écrit généralement sous la forme d'une *fraction* (d'un *rapport* ou d'un *quotient*) $\frac{a}{b}$ qu'on représente encore, par souci d'économie, à l'aide d'une barre oblique a/b. Donc, la fraction a/b qui n'a un sens que pour $b \neq 0$, est solution unique de l'équation $bx = a$. Les règles de calcul des fractions sont les suivantes :

$$\begin{aligned} &\frac{a}{b} = \frac{c}{d} \Leftrightarrow ad = bc, && b, d \neq 0, \\ &\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}, && b, d \neq 0, \\ &-\frac{a}{b} = \frac{-a}{b} = \frac{a}{-b} && b \neq 0, \\ &\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}, && b, d \neq 0. \\ &\left(\frac{a}{b}\right)^{-1} = \frac{b}{a} && a, b \neq 0. \end{aligned} \tag{10}$$

Ce sont des règles usuelles « d'école » qu'il ne s'agit pas d'apprendre par cœur, mais qu'il faut déduire des axiomes du corps commutatif, ce qui ne présente d'ailleurs aucune difficulté. Voilà les raisonnements qui suffisent pour obtenir la deuxième des règles (10). Soient $x = a/b$ et $y = c/d$ solutions des équations $bx = a$ et $dy = c$. Il en résulte que $dbx = da$, $bdy = bc \Rightarrow bd(x + y) = da + bc \Rightarrow t = x + y = (da + bc)/bd$ est solution unique de l'équation $bdt = da + bc$.

On appelle *sous-corps* F d'un corps P tout sous-anneau de P qui est lui-même un corps. Par exemple, le corps des nombres rationnels $\mathbb{Q}$ est un sous-corps du corps des nombres réels $\mathbb{R}$.

Dans le cas où $F \subset P$, on dit que le corps P est une *extension* de son sous-corps F. De la définition du sous-corps il résulte que l'élément zéro et l'élément unité du corps P appartiennent également à F et sont pour F élément zéro et élément unité. Soit F_1 l'intersection de tous les sous-corps de P contenant F et un certain élément $a \in P$, tel que $a \notin F$. Alors F_1 est un corps minimal qui contient l'ensemble $\{F, a\}$ (même raisonnement que celui développé pour les groupes au § 2, nº 2). On dit que l'extension F_1 du corps F est obtenue par *adjonction* à F de l'élément a, et on reflète ce fait par la notation $F_1 = F(a)$. De même, on peut parler d'un sous-corps $F_1 = F(a_1, \ldots, a_n)$ du corps P, obtenu par adjonction à F de n éléments $a_1, \ldots, a_n$ du corps P.

Une petite vérification montre que $\mathbb{Q}(\sqrt{2})$ coïncide avec l'ensemble des nombres $a+b\sqrt{2}$, $a, b \in \mathbb{Q}$, car $(\sqrt{2})^2 = 2$ et $1/(a+b\sqrt{2}) = (a/(a^2-2b^2)) - (b/(a^2-2b^2))\sqrt{2}$ pour $a+b\sqrt{2} \neq 0$. La même remarque est valable pour $\mathbb{Q}(\sqrt{3})$, $\mathbb{Q}(\sqrt{5})$, etc.

On dit que deux corps P et P' sont *isomorphes*, s'ils le sont en tant qu'anneaux. Il résulte immédiatement de la définition que $f(0) = 0$ et $f(1) = 1'$ pour toute application isomorphe f. Parler des homomorphismes des corps n'a aucun sens, car $\operatorname{Ker} f \neq 0 \Rightarrow f(a) = 0$, $a \neq 0 \Rightarrow f(1) = f(aa^{-1}) = f(a) f(a^{-1}) = 0 \cdot f(a^{-1}) = 0 \Rightarrow f(b) = f(1 \cdot b) = f(1) f(b) = 0 \cdot f(b) = 0, \forall\, b \Rightarrow \operatorname{Ker} f = P$. En revanche, les *automorphismes*, c'est-à-dire les applications isomorphes d'un corps P sur lui-même, sont liés aux propriétés les plus profondes des corps et fournissent un instrument puissant pour l'étude de ces propriétés dans le cadre de la *théorie de Galois*.

La notion d'extension des corps est à l'unisson du désir éternel des hommes à élargir l'arsénal des nombres utilisés. Un processus assez lent qu'on peut représenter conventionnellement par le diagramme: {un} $\rightsquigarrow$ {un et un font deux} $\rightsquigarrow \mathbb{N} \rightsquigarrow \{\mathbb{N}, 0\} \rightsquigarrow \mathbb{Z} \rightsquigarrow \mathbb{Q} \rightsquigarrow \mathbb{Q}(\sqrt{2}) \rightsquigarrow \mathbb{R}$ et qui a continué à se dérouler jusqu'à présent, a conduit à un réseau extrêmement développé de corps qui diffèrent beaucoup des corps numériques habituels. Toutes les étapes

de ce processus n'ont pas été purement algébriques. Par exemple, le passage des nombres rationnels aux nombres *réels*, basé sur les notions de continuité et de plénitude (existence de limites pour les suites de Cauchy), est étudié jusqu'à présent dans les cours d'analyse mathématique. En même temps, la construction tout à fait analogue des corps des nombres p-adiques que nous n'analysons pas ici, et l'analyse p-adique moderne qui s'est développée sur la base de cette construction, sont de belles œuvres de trois branches des mathématiques : de la théorie des nombres, de l'algèbre et de l'analyse.

6. Caractéristique d'un corps commutatif.— Au nº 2, nous avons construit un anneau fini des classes résiduelles Z_m ayant les éléments

$$\overline{0},\ \overline{1},\ \overline{2},\ \ldots,\ \overline{m-1}$$

et muni des opérations d'addition $\overline{k}+\overline{l}=\overline{k+l}$, et de multiplication $\overline{k}\cdot\overline{l}=\overline{kl}$ (nous renonçons à l'emploi des signes $\oplus$ et $\odot$). Si $m = st$, $s > 1$, $t > 1$, on a $\overline{s}\cdot\overline{t}=\overline{m}=\overline{0}$, c'est-à-dire $\overline{s}$ et $\overline{t}$ sont des diviseurs de zéro dans Z_m.

Soit maintenant $m = p$ un nombre premier. On affirme que Z_p est un corps (à p éléments). Pour $p = 2$ et 3 cette assertion découle immédiatement des tables de multiplication écrites au nº 2. Dans le cas général, il suffit d'établir pour tout $\overline{s} \in Z_p^*$ l'existence d'un élément inverse $\overline{s'}$ (les entiers s et s' ne doivent pas manifestement être divisibles par p).

Considérons les éléments

$$\overline{s},\ \overline{2s},\ \ldots,\ \overline{(p-1)s}. \tag{11}$$

Ils sont tous non nuls, car $s \not\equiv 0 \pmod{p} \Rightarrow ks \not\equiv 0 \pmod{p}$ pour $k = 1, 2, \ldots, p-1$. On utilise ici le fait que p est un nombre premier. Pour la même raison, les éléments (11) sont tous distincts : $\overline{ks}=\overline{ls}$, $k < l$, entraînerait $\overline{(l-k)s}=\overline{0}$, ce qui est faux. Ainsi, la suite des éléments (11) coïncide avec celle des éléments

$$\overline{1},\ \overline{2},\ \ldots,\ \overline{p-1}$$

permutés d'une façon quelconque.

En particulier, il existe un élément s', $1 \leqslant s' \leqslant p-1$, tel que l'on ait $\overline{s's}=\overline{1}$. Or, cela signifie justement que $\overline{s'}\cdot\overline{s}=\overline{1}$, c'est-à-dire que $\overline{s'}$ est l'élément inverse de $\overline{s}$. Nous avons ainsi démontré le théorème suivant :

THÉORÈME 5. — *Un anneau des classes résiduelles Z_m est un corps si, et seulement si, $m = p$ est un nombre premier.* ∎

COROLLAIRE (théorème de Fermat). — *Pour tout module premier p, et tout entier m non divisible par p, on a la congruence*

$$m^{p-1} \equiv 1 \pmod{p}.$$

Démonstration. — Le groupe multiplicatif Z_p^* est d'ordre $p - 1$. D'après le théorème de Lagrange (voir § 3), $p - 1$ est divisible par l'ordre de tout élément de Z_p^*. Ainsi $\overline{1} = (\overline{m})^{p-1} = \overline{m^{p-1}}$, c'est-à-dire $\overline{m^{p-1} - 1} = \overline{0}$. ■

Il n'est pas difficile de démontrer le théorème de Fermat directement à l'aide de la théorie des congruences, en multipliant tous les éléments de la suite (11).

Les corps Z_2, Z_3, Z_5, . . ., si peu ressemblants aux corps connus $\mathbb{Q}$, $\mathbb{Q}(\sqrt{2})$, $\mathbb{R}$, ont pris dans l'hiérarchie algébrique une place qui est comparable, d'après son importance, à celle attribuée depuis longtemps au corps $\mathbb{Q}$. L'explication en est la suivante. Soit P un corps. Comme nous l'avons déjà indiqué, l'intersection $\bigcap_i P_i$ de toute famille de sous-corps $\{P_i \mid i \in I\}$ est un sous-corps de P.

Définition. — *Un corps est dit premier s'il ne possède aucun sous-corps strict.*

Théorème 6. — *Tout corps commutatif P contient un corps premier P_0 et un seul. Ce corps premier est isomorphe soit à $\mathbb{Q}$, soit à Z_p pour un certain p.*

Démonstration. — Il est évident que l'intersection P_0 de tous les sous-corps du corps P ne contient pas de sous-corps stricts et est l'unique sous-corps de P vérifiant cette propriété.

Le corps P_0 contient avec l'élément unité 1 tous ses multiples $n \cdot 1 = 1 + \ldots + 1$. Il résulte des propriétés générales des opérations d'addition et de multiplication des éléments dans les anneaux (voir la fin du nº 1) que

$$s \cdot 1 + t \cdot 1 = (s + t) \cdot 1, \quad (s \cdot 1)(t \cdot 1) = (st) \cdot 1; \quad s, t \in \mathbb{Z}. \qquad (12)$$

C'est pourquoi, l'application f de l'anneau $\mathbb{Z}$ dans P, définie par la loi $f(n) = n1$, est un homomorphisme dont le noyau, qui est un idéal de $\mathbb{Z}$, est de la forme $\operatorname{Ker} f = m\mathbb{Z}$. Si $m = 0$, f est un isomorphisme, et les fractions $(s \cdot 1)/(t \cdot 1)$ qui ont un sens dans P (puisque P est un corps commutatif) forment un corps P_0 isomorphe à $\mathbb{Q}$. Il sera justement le sous-corps premier de P.

Si $m > 0$, il est évident que l'application f^* définie par la loi

$$f^* : \overline{k} = \{k\}_m \to f(k),$$

sera un plongement isomorphe $Z_m \to P$, ce qui est possible si, et seulement si, $m = p$ est un nombre premier, car pour m non premier Z_m contient les diviseurs de zéro. Donc, $f^*(Z_p)$ est un sous-corps premier de P. ■

Définition. — *On dit qu'un corps commutatif P est de caractéristique nulle si son sous-corps premier P_0 est isomorphe à $\mathbb{Q}$; le corps P*

est de caractéristique un nombre premier p (ou de caractéristique finie) si $P_0 \cong Z_p$. *On écrit respectivement* car $P = 0$ ou car $P = p > 0$.

Pour désigner un corps « abstrait » à p éléments, au lieu de Z_p on utilise généralement $\mathbb{F}_p$ ou GF (p) (*Galois Field-corps de Galois*). Il faut avoir en vue qu'il existe un corps commutatif fini GF (q) à $q = p^n$ éléments, où p est un nombre premier, et n un entier positif quelconque. Nous reviendrons sur cette question intéressante au chapitre 9, en nous contenant d'indiquer ici un exemple de corps commutatif à quatre éléments $\{0, 1, \alpha, \beta\}$:

GF (4):

+	0	1	α	β
0	0	1	α	β
1	1	0	β	α
α	α	β	0	1
β	β	α	1	0

·	0	1	α	β
0	0	0	0	0
1	0	1	α	β
α	0	α	β	1
β	0	β	1	α

La nature des éléments α et β ne nous intéresse pas pour l'instant. Nous recommandons au lecteur de s'assurer que la loi de distributivité est vérifiée.

La caractéristique nulle est appelée parfois caractéristique infinie, conformément à son interprétation en tant qu'ordre de l'élément 1 dans le groupe additif du corps P. De même, la caractéristique finie p est l'ordre général de tout élément non nul du groupe additif:

$$px = x + \ldots + x = 1 \cdot x + \ldots + 1 \cdot x = \\ = (1 + \ldots + 1)\, x = (p \cdot 1)\, x = 0.$$

7. **Remarque sur les systèmes linéaires.**— Il est déjà temps d'envelopper d'un regard, en pensée, la théorie des systèmes d'équations linéaires que nous avons exposée au cours des chapitres précédents, et la théorie des déterminants qu'elle a fait naître. Comme coefficients d'équations linéaires et éléments de matrices nous avons utilisé des nombres rationnels ou réels dont le caractère spécifique ne nous importait pas. Rien ne nous empêche maintenant de prendre au lieu de ces nombres des éléments d'un corps commutatif arbitraire P. Dans ce cas, les résultats obtenus doivent être énoncés en termes de corps P: les composantes de la solution d'un système linéaire et les valeurs de la fonction det appartiendront au corps P. La méthode de Gauss de résolution des systèmes d'équations linéaires, la théorie des déterminants, les formules de Cramer restent valables (sans modifications importantes) pour un corps commutatif arbitraire P.

EXEMPLE 1. — Soient donnés un système homogène d'équations linéaires $AX = 0$ de matrice carrée

$$A = (a_{ij}) = \left\| \begin{matrix} 1 & 2 & 3 & 4 \\ -10 & 13 & 14 & 1 \\ 12 & -9 & 14 & 15 \\ 12 & 13 & -8 & 15 \end{matrix} \right\|$$

et une colonne d'inconnues $X = [x_1, x_2, x_3, x_4]$. Le calcul direct montre que $\det A = 2^3 \cdot 11^3$. Par conséquent, pour $a_{ij}, x_k \in P$, où P est un corps commutatif arbitraire de caractéristique nulle ou de caractéristique $p \neq 2, 11$ (dans ce cas les entiers 1, 2, 3, 4, —10, . . ., 15 sont remplacés par les classes résiduelles correspondantes),[1] notre système est un système déterminé qui n'admet qu'une solution triviale $X = 0$.

Si car $P = 2$ (disons, $P = Z_2$), la congruence

$$\left\| \begin{matrix} 1 & 2 & 3 & 4 \\ -10 & 13 & 14 & 15 \\ 12 & -9 & 14 & 15 \\ 12 & 13 & -8 & 15 \end{matrix} \right\| \equiv \left\| \begin{matrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{matrix} \right\| \pmod{2}$$

permet de conclure que le rang du système est égal à deux et que le système admet deux solutions indépendantes $X_1 = [1, 0, 1, 0]$, $X_2 = [0, 1, 0, 1]$. Pour éviter tout malentendu, il faudrait écrire $X_1 = [\bar{1}, \bar{0}, \bar{1}, \bar{0}]$, $X_2 = [\bar{0}, \bar{1}, \bar{0}, \bar{1}]$, mais on est assez préparé pour se contenter de la notation simplifiée.

Si car $P = 11$, on déduit de la congruence

$$\left\| \begin{matrix} 1 & 2 & 3 & 4 \\ -10 & 13 & 14 & 15 \\ 12 & -9 & 14 & 15 \\ 12 & 13 & -8 & 15 \end{matrix} \right\| \equiv \left\| \begin{matrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{matrix} \right\| \pmod{11}$$

que le système possède trois solutions indépendantes:

$$X_1 = [9, 1, 0, 0], \quad X_2 = [8, 0, 1, 0], \quad X_3 = [7, 0, 0, 1].$$

Comme nous le voyons, la réponse dépend essentiellement du corps P considéré. En revanche, l'analyse du système ne diffère en rien de l'analyse ordinaire. Donc, l'un des avantages offerts par le passage de $\mathbb{R}$ et $\mathbb{Q}$ à un corps arbitraire est qu'il permet d'éviter les doubles raisonnements similaires. Mais il existe d'autres raisons plus profondes qui militent en faveur de ce passage.

En parlant du groupe linéaire complet, nous l'avons considéré jusqu'ici comme groupe de toutes les matrices régulières à coefficients dans $\mathbb{Q}$ ou $\mathbb{R}$. L'ensemble des matrices carrées d'ordre n à coefficients dans un corps arbitraire P forme un anneau des matrices $M_n(P)$, et le sous-ensemble de toutes les matrices régulières $A \in M_n(P)$ (des matrices de $\det A \neq 0$) conduit à la notion de groupe linéaire complet $GL(n, P)$ sur P. Faisant varier le corps P, en posant $P = \mathbb{F}_p$ par exemple, on peut obtenir de façon naturelle toute une série de groupes importants (voir chap. 7).

Les corps de type $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{Q}(\sqrt{2})$, etc., sont généralement appelés *corps numériques*. Le corps $\mathbb{F}_p$ est un exemple de corps non numérique: il serait incorrect d'appeler ses éléments nombres pour cette seule raison qu'ils sont souvent identifiés avec les éléments de l'ensemble $\{0, 1, \ldots, p-1\}$.

Au chapitre 1, § 2, il a été posé un problème (sous le numéro 3) concernant l'emploi des corps finis dans la théorie du codage. Nous donnons ci-dessous un petit exemple à ce sujet.

EXEMPLE 2. — Pour transmettre l'appel МИРУ МИР (PAIX AU MONDE) il suffit en principe de répéter, un nombre de fois correspondant et dans l'ordre convenable, quatre messages élémentaires $M = (0, 0)$, И $= (1, 0)$, Р $= (0, 1)$, У $= (1, 1)$, interprétés comme vecteurs lignes d'un espace vectoriel de dimension deux $\mathbb{F}_2^2$ sur le corps $\mathbb{F}_2 \cong Z_2 = \{0, 1\}$ à deux éléments. Le bruit intervenant dans le canal au cours d'une transmission fait naître des perturbations (le remplacement du symbole 0 par 1 ou de 1 par 0), de sorte qu'un message РИМУ РИМ (ROME à ROME) par exemple peut arriver au point de réception. D'après le théorème fondamental de Shannon, l'influence des parasites peut être éliminée par une augmentation de la longueur des messages élémentaires (c'est-à-dire en diminuant la vitesse de transmission). Supposons connu des conditions de transmission que chaque message élémentaire de longueur cinq a tout au plus une erreur. On prend alors dans l'espace vectoriel $S = \mathbb{F}_2^5$ le sous-ensemble $S_0 = \{$М $= (0, 0, 1, 1, 0)$, И $= (1, 0, 0, 1, 1)$, Р $= (0, 1, 1, 0, 1)$, У $=$ $= (1, 1, 0, 0, 0)\}$ appelé sous-ensemble de *vecteurs codes*. Le tableau

Vecteurs codes	00110	10011	01101	11000
Vecteurs obtenus à partir des vecteurs codes par suite d'une distorsion	00010 00100 00111 01110 10110	00011 10001 10010 10111 11011	00101 01001 01100 01111 11101	01000 10000 11100 11001 11010

montre que les ensembles de vecteurs déformés des différentes colonnes sont disjoints, ce qui signifie que le décodage correct est possible, c'est-à-dire le destinataire reçoit le message tel qu'il a été émis.

Nous avons ainsi obtenu un *code S_0 corrigeant une erreur*. En passant à des espaces $\mathbb{F}_2^n$ de dimension n suffisamment grande, on peut construire un code analogue permettant de transmettre fidèlement tout l'alphabet, c'est-à-dire n'importe quel texte. Pour que le décodage ne soit pas ramené à un très long et lent triage, on est amené à choisir S_0 d'une façon spéciale. A cet effet, on a mis au point de nombreuses méthodes, y compris purement algébriques, basées sur l'emploi des corps finis $\mathbb{F}_q$.

EXERCICES

1. En développant l'idée de l'exemple 2 du § 1, montrer que l'ensemble $\mathcal{P}(\Omega)$ muni des opérations

$$A + B = (A \cup B) \setminus (A \cap B), \quad AB = A \cap B; \quad A, B \in \Omega,$$

est un anneau unitaire, et que tous les éléments de son groupe additif sont d'ordre 2.

2. Etablir la commutativité d'un anneau arbitraire dont tout élément x satisfait à l'équation $x^2 = x$. Est-ce vrai pour $x^3 = x$?

3. Les corps $\mathbb{Q}(\sqrt{2})$, $\mathbb{Q}(\sqrt{5})$ sont-ils isomorphes?

4. Les éléments non inversibles de l'anneau: 1) Z_{16}; 2) Z_{24} forment-ils un idéal?

12—0877

5. Montrer que l'image épimorphe d'un anneau commutatif est un anneau commutatif.

6. Démontrer que si K est un anneau unitaire et L un idéal, l'anneau quotient K/L a, lui aussi, un élément unité.

7. Montrer que tout anneau intègre fini K est un corps.

8. Soient p un nombre premier et K un anneau unitaire commutatif tel que $px = 0$ pour tout $x \in K$. Montrer que dans ces conditions

$$(x+y)^{p^m} = x^{p^m} + y^{p^m}, \quad m = 1, 2, \ldots$$

(Indication. Utiliser la récurrence sur m et le fait que le coefficient binomial $\binom{p}{k}$, $0 < k < p$, est divisible par p.)

9. Démontrer qu'un anneau K à cinq éléments est isomorphe à Z_5 ou est un anneau à multiplication nulle.

10. L'ensemble $T = \left\{ \begin{Vmatrix} a & c \\ 0 & b \end{Vmatrix} \mid a, b \in \mathbb{Z} \right\}$ des matrices triangulaires supérieures forme un sous-anneau de $M_2(\mathbb{Z})$. S'en assurer et donner la description des idéaux de l'anneau T.

11. Un élément $x \neq 0$ d'un anneau K est dit *nilpotent* si $x^n = 0$ pour un $n \in \mathbb{N}$. Montrer que:

(i) l'élément $1 - x$ est nilpotent dans tout anneau unitaire si x est inversible;

(ii) l'anneau $Z_m = \mathbb{Z}/m\mathbb{Z}$ contient des éléments nilpotents si, et seulement si, m est divisible par le carré d'un entier naturel >1.

12. Démontrer qu'un anneau unitaire K de puissance infinie $|K|$ ne peut pas contenir un nombre fini $n \geqslant 1$ d'éléments non inversibles $\neq 0$. (Indication. Raisonner par l'absurde. Soit $N = \{a_1, \ldots, a_n\}$ l'ensemble de tous les éléments non inversibles $\neq 0$ de l'anneau K. L'application ρ_x: $a_i \mapsto xa_i$ est une bijection $N \to N$ pour tout $x \in K \setminus (N \cup \{0\})$. Le noyau $\operatorname{Ker} \rho$ de l'application ρ: $x \mapsto \rho_x$ est infini.)

13. Soit K un anneau associatif arbitraire possédant un élément unité 1 et soient a, b ses éléments. Montrer que

$$(1 - ab)\, c = 1 = c\,(1 - ab) \Rightarrow (1 - ba)\, d = 1 = d\,(1 - ba),$$

où $d = 1 + bca$, c'est-à-dire si $1 - ab$ est inversible dans K, il en est de même de $1 - ba$. Quel est l'élément $1 + adb$?

14. Montrer que les matrices $\begin{Vmatrix} a & b \\ -b & a \end{Vmatrix}$, avec $a, b \in Z_3$, forment un corps commutatif à 9 éléments, dont le groupe multiplicatif est un groupe cyclique d'ordre 8.

15. Le code S_0 (voir exemple 2 à la fin du paragraphe) est-il capable de corriger deux erreurs?

CHAPITRE 5

NOMBRES COMPLEXES ET POLYNÔMES

Le présent chapitre est consacré à des systèmes algébriques tout à fait concrets, qui sont partiellement connus du cours de mathématiques de l'enseignement secondaire, mais qui méritent d'être étudiés avec plus de détails. Le point de vue élaboré au cours du chapitre précédent nous permettra de voir sous un jour nouveau le « champ d'action » traditionnel de l'algèbre des siècles passés. En même temps, l'étude des polynômes, par exemple, rendra plus compréhensibles et tangibles des problèmes tels que l'extension des anneaux et l'unicité de la décomposition en facteurs premiers dans les anneaux intègres (dans les domaines d'intégrité).

§ 1. Corps des nombres complexes

L'histoire des mathématiques connaît une lutte opiniâtre entre les partisans et les adversaires des nombres « imaginaires » dont l'origine est due à l'équation algébrique

$$x^2 + 1 = 0. \tag{1}$$

On peut certes prendre une position simpliste et se contenter d'une écriture formelle des solutions de l'équation (1) sous la forme $\pm\sqrt{-1}$. Or, ceci pouvait être fait sans peine aussi bien dans des temps plus éloignés ; il ne restait qu'à donner un sens à cette écriture. Nous allons résoudre ce problème à des niveaux différents. Commençons par donner quelques considérations heuristiques.

1. Construction auxiliaire.— Proposons-nous de construire une extension du corps $\mathbb{R}$ des nombres réels où l'équation (1) possède une solution. Considérons à titre de modèle d'une telle extension l'ensemble P de toutes les matrices carrées de type

$$\begin{Vmatrix} a & b \\ -b & a \end{Vmatrix} \in M_2(\mathbb{R}). \tag{2}$$

On affirme que P est un corps commutatif (comparer avec chap. 4, § 4, exercice 14).

En effet, P contient le zéro 0 et l'élément unité E de l'anneau $M_2(\mathbb{R})$.

On a aussi les relations

$$\begin{Vmatrix} a & b \\ -b & a \end{Vmatrix} + \begin{Vmatrix} c & d \\ -d & c \end{Vmatrix} = \begin{Vmatrix} a+c & b+d \\ -(b+d) & a+c \end{Vmatrix},$$
$$-\begin{Vmatrix} a & b \\ -b & a \end{Vmatrix} = \begin{Vmatrix} -a & -b \\ -(-b) & -a \end{Vmatrix}, \qquad (3)$$
$$\begin{Vmatrix} a & b \\ -b & a \end{Vmatrix} \begin{Vmatrix} c & d \\ -d & c \end{Vmatrix} = \begin{Vmatrix} ac-bd & ad+bc \\ -(ad+bc) & ac-bd \end{Vmatrix},$$

d'où il découle que P est stable pour l'addition et la multiplication. L'associativité de ces opérations résulte de leur associativité dans M_2. Il en est de même bien attendu, pour les lois de distributivité et de commutativité de l'addition. Ainsi, P est un sous-anneau de M_2. Il reste à démontrer que toute matrice de P de déterminant $\begin{vmatrix} a & b \\ -b & a \end{vmatrix} = a^2 + b^2 \neq 0$ possède dans P une matrice inverse (la commutativité de P découle des formules (3)). On obtient en effet

$$\begin{Vmatrix} a & b \\ -b & a \end{Vmatrix}^{-1} = \begin{Vmatrix} c & d \\ -d & c \end{Vmatrix}, \quad \text{avec} \quad c = \frac{a}{a^2+b^2}, \quad d = \frac{-b}{a^2+b^2}, \qquad (4)$$

soit directement par la formule donnant les coefficients de la matrice inverse (voir chap. 3, § 3, théorème 1), soit en résolvant le système linéaire

$$ax - by = 1,$$
$$bx + ay = 0,$$

provenant de la condition

$$\begin{Vmatrix} a & b \\ -b & a \end{Vmatrix} \begin{Vmatrix} x & y \\ -y & x \end{Vmatrix} = \begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix}.$$

En se servant de la règle de multiplication des matrices par des nombres (voir chap. 2, § 3, règle (5)), on peut écrire tout élément du corps P sous forme de

$$\begin{Vmatrix} a & b \\ -b & a \end{Vmatrix} = aE + bJ, \quad \text{où} \quad a, b \in \mathbb{R}, \quad J = \begin{Vmatrix} 0 & 1 \\ -1 & 0 \end{Vmatrix}. \qquad (5)$$

L'ensemble $\{aE \mid a \in \mathbb{R}\} \cong \mathbb{R}$ est un sous-corps de P, et la relation

$$J^2 + E = 0$$

montre que l'élément $J \in P$ est, « à un isomorphisme près », solution de l'équation (1). Donc, il ne s'agit ici d'aucun mystère de la « quantité imaginaire J ».

Cependant, ce n'est pas le corps P qu'on appelle corps des nombres complexes, mais un certain être mathématique qui lui est isomorphe et dont les éléments se représentent par des points d'un plan. Le désir d'avoir une réalisation géométrique du corps P est tout à fait naturel si l'on se rappelle que le corps $\mathbb{R}$ est pour nous inséparable de la « droite réelle » dont un point fixe représente 0 et un repère choisi définit la position du nombre 1.

2. Plan des nombres complexes.— Ainsi, nous voulons construire un corps commutatif $\mathbb{C}$ dont les éléments sont des points du plan $\mathbb{R}^2$, alors que l'addition et la multiplication dans $\mathbb{C}$ obéissent à toutes les règles concernant les opérations dans un corps commutatif et résolvent notre problème. Rapportons le plan cartésien à un repère orthogonal, avec x, l'axe des abscisses, et y, l'axe des ordonnées. Notons (a, b) le point d'abscisse a et d'ordonnée b, et définissons pour les points (a, b) et (c, d) la somme et le produit :

$$\begin{aligned} (a, b) + (c, d) &= (a + c, b + d), \\ (a, b)(c, d) &= (ac - bd, ad + bc) \end{aligned} \tag{6}$$

(l'emploi des mêmes signes $+$ et $\cdot$ que dans le corps $\mathbb{R}$ ne doit mener à aucune confusion). Une vérification directe, mais assez fastidieuse, prouverait que les opérations ainsi définies confèrent à l'ensemble des couples (points du plan) une structure de corps commutatif vérifiant les propriétés exigées. Heureusement, une telle vérification n'est pas nécessaire. La correspondance

$$(a, b) \mapsto \begin{Vmatrix} a & b \\ -b & a \end{Vmatrix}$$

qui aux points du plan $\mathbb{C}$ associe les éléments du corps commutatif P, et un coup d'œil rapide, jeté sur les formules (3) et (6), nous prouvent qu'il y a ici un isomorphisme et que l'ensemble $\mathbb{C}$ est donc un corps commutatif. C'est à lui qu'on donne habituellement le nom de corps des nombres complexes. Eu égard à la réalisation géométrique de ce corps, $\mathbb{C}$ est encore appelé *plan des nombres complexes* (ou encore *plan complexe*, ce qui arrive plus souvent, quoique ce soit une appellation un peu biunivoque).

L'axe des abscisses que nous avons choisi, c'est-à-dire l'ensemble des points $(a, 0)$, ne diffère en rien, par ses propriétés, de la droite réelle et nous posons $(a, 0) = a$. Le zéro $(0, 0)$ et l'unité $(1, 0)$ du corps deviennent dans ces conditions des nombres réels ordinaires. Pour le point $(0, 1)$ situé sur l'axe des ordonnées, on introduit la désignation traditionnelle i de l'« unité imaginaire », qui est racine de l'équation (1) : $i^2 = (0, 1)(0, 1) = (-1, 0) = -1$. Un nombre complexe arbitraire $z = (x, y)$ s'écrira maintenant sous la forme habituelle

$$z = x + iy, \quad x, y \in \mathbb{R}, \tag{7}$$

qui se rapproche sensiblement de la forme (5) des éléments du corps P. Remarquons que $\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$. Aussi, $\mathbb{C}$ est-il un corps de caractéristique nulle (voir chap. 4, § 4, nº 6).

3. Interprétation géométrique des opérations sur les nombres complexes.— L'axe des abscisses du plan $\mathbb{C}$ est généralement appelé *axe réel*, et l'axe des ordonnées, *axe imaginaire*, alors que les nombres iy situés sur cet axe sont dits *nombres imaginaires purs* bien que le terme « *imaginaire* » ait perdu son sens primitif. Respectivement, dans l'expression (7), le nombre x s'appelle *partie réelle* du nombre z, et iy, *partie imaginaire* de z. Considérons une application

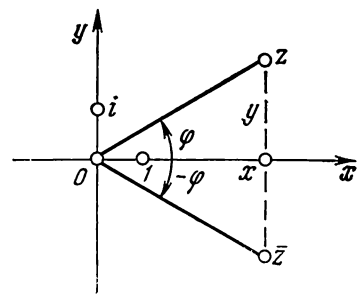


Fig. 15.

qui à chaque nombre complexe $z = x + iy$ associe un nombre complexe $\bar{z} = x - iy$ appelé *conjugué de z* (cette opération s'appelle *conjugaison imaginaire*). Géométriquement, cette application se ramène à la transformation de symétrie orthogonale du plan $\mathbb{C}$ par rapport à l'axe réel (voir fig. 15). On peut énoncer le théorème bien instructif suivant:

THÉORÈME 1. — *L'application $z \mapsto \bar{z}$ est un automorphisme d'ordre* 2 *du corps $\mathbb{C}$ qui laisse fixes tous les nombres réels. La somme et le produit des nombres complexes conjugués sont des nombres réels.*

DÉMONSTRATION. — L'affirmation que $\bar{x} = x$, $x \in \mathbb{R}$, est évidente, vu la définition du nombre complexe conjugué.

En particulier, on a $\bar{0} = 0$ et $\bar{1} = 1$. L'assertion concernant l'ordre est, elle aussi, évidente: $\overline{(\bar{z})} = z$. Il ne nous reste qu'à vérifier les relations

$$\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2, \quad \overline{z_1 z_2} = \bar{z}_1 \cdot \bar{z}_2. \tag{8}$$

Or, elles résultent directement des formules (6) lorsque celles-ci sont écrites sous la forme

$$\begin{aligned} (x_1 + iy_1) + (x_2 + iy_2) &= (x_1 + x_2) + i\,(y_1 + y_2), \\ (x_1 + iy_1)\cdot(x_2 + iy_2) &= (x_1x_2 - y_1y_2) + i\,(x_1y_2 + x_2y_1). \end{aligned} \tag{9}$$

Un cas particulier des formules (9) est l'assertion relative à la somme et au produit de deux nombres complexes conjugués $z = x + iy$ et $\bar{z}$: $z + \bar{z} = 2x$; $z\bar{z} = x^2 + y^2$. ■

REMARQUE. — L'automorphisme $z \mapsto \bar{z}$ se distingue de nombreux autres automorphismes du corps $\mathbb{C}$ par le fait qu'il est l'unique automorphisme continu (qui transforme deux points voisins du plan $\mathbb{C}$ en des points voisins). Nous ne précisons et ne démontrons pas cette proposition.

On appelle *module* (ou *valeur absolue*) d'un nombre complexe $z = x + iy$ le nombre réel non négatif $|z| = \sqrt{z\bar{z}} = \sqrt{x^2 + y^2}$. On sait que la position du point z sur le plan est entièrement déterminée par la donnée de ses coordonnées polaires: distance $r = |z|$ du point z à l'origine des coordonnées et angle φ entre la direction positive de l'axe des abscisses et le rayon vecteur de z (voir fig. 15). L'angle φ s'appelle *argument* du nombre z et se note $\arg z = \varphi$. En vertu de la définition même, $\arg z$ peut prendre toutes les valeurs positives et négatives, mais pour un r donné, les angles qui diffèrent l'un de l'autre d'un multiple entier de 2π, correspondent à un seul et même nombre. L'argument n'est pas défini pour le nombre 0 de module $|0| = 0$. Etant appliquées aux nombres complexes, les relations « être supérieur à » ou « être inférieur à » n'ont pas de sens, autrement dit les nombres complexes ne peuvent pas être réunis par un signe d'inégalité: à la différence des nombres réels dont l'argument ne peut prendre que deux valeurs principales: 0 (pour les nombres positifs) et π (pour les nombres négatifs), *les nombres complexes ne sont pas ordonnés.*

Les coordonnées polaires r et φ déterminent x et y à l'aide des formules bien connues

$$x = r \cos \varphi, \quad y = r \sin \varphi, \quad z = r (\cos \varphi + i \sin \varphi). \tag{10}$$

C'est la *forme* dite *trigonométrique* du nombre z.

L'opération d'addition des nombres complexes z, z' s'exprime de façon très simple en coordonnées cartésiennes, à savoir selon la règle du parallélogramme ou, ce qui revient au même, selon la règle d'addition des segments orientés (vecteurs) qui partent de l'origine des coordonnées et qui correspondent aux nombres z, z' (voir fig. 16). En comparant les côtés du triangle de sommets aux points 0, z et $z + z'$ (et en identifiant les valeurs absolues à des longueurs géométriques correspondantes) on obtient de la même figure une inégalité importante

$$|z + z'| \leqslant |z| + |z'| \tag{11}$$

connue sous le nom d'inégalité triangulaire. Remarquons que l'inégalité (11) qui peut s'écrire sous une forme plus générale

$$|z| - |z'| \leqslant |z \pm z'| \leqslant |z| + |z'|$$

est tout à fait analogue à l'inégalité correspondante valable pour les nombres réels.

Le produit des nombres complexes s'exprime commodément en coordonnées polaires.

THÉORÈME 2. — *Le module du produit de deux nombres complexes z, z' est égal au produit de leurs modules, et l'argument, à la somme de leurs arguments:*

$$|zz'| = |z| \cdot |z'|, \quad \arg zz' = \arg z + \arg z'. \tag{12}$$

D'une manière analogue, $|z/z'| = |z|/|z'|$, $\arg z/z' = \arg z - \arg z'$.

DÉMONSTRATION. — Soient

$$z = r(\cos\varphi + i\sin\varphi), \quad z' = r'(\cos\varphi' + i\sin\varphi')$$

deux nombres complexes présentés sous la forme trigonométrique

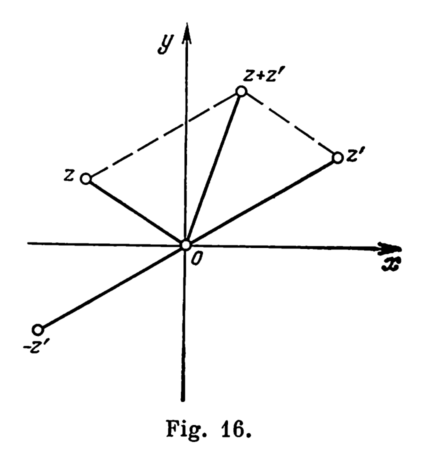


Fig. 16.

(10). Par multiplication directe ou en se servant de la formule (9) on obtient

$$zz' = rr'[(\cos\varphi\cos\varphi' - \sin\varphi\sin\varphi') + i(\cos\varphi\sin\varphi' + \sin\varphi\cos\varphi')],$$

d'où en appliquant des formules connues, il vient la forme trigonométrique du nombre zz':

$$zz' = |z| \cdot |z'| \cdot [\cos(\varphi + \varphi') + i\sin(\varphi + \varphi')].$$

Si, ensuite, $z'' = z/z'$, alors $z = z'z''$. Aussi, en utilisant les formules (12) démontrées pour le produit $z'z''$, peut-on en déduire les formules pour le quotient z/z'. ■

On a, en particulier, $z^{-1} = |z|^{-1} [\cos(-\varphi) + i \sin(-\varphi)]$. Pour obtenir z^{-1} sur le plan complexe (voir fig. 17), il faut donc appliquer à z d'abord une inversion par rapport à un cercle de rayon unité

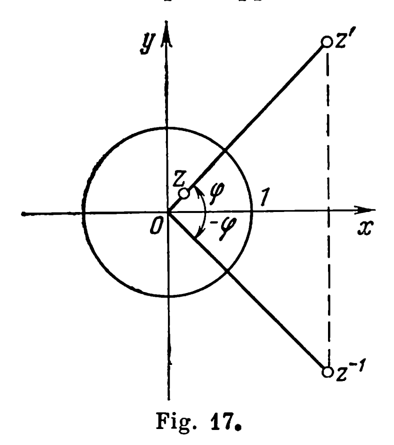


Fig. 17.

et de centre en 0 (ceci donne le point z') et puis une réflexion par rapport à l'axe réel (ou l'automorphisme $z' \mapsto \bar{z}'$).

En réalité, les propositions relatives au module du produit et au module de la somme se déduisent aisément du théorème 1 sans avoir recours à l'intuition géométrique. En effet, premièrement

$$|zz'|^2 = zz'\overline{zz'} = zz'\bar{z}\bar{z}' = z\bar{z}\cdot z'\bar{z}' = |z|^2 |z'|^2,$$

d'où $|zz'| = |z|\cdot|z'|$. Puis, en remarquant que $|z| = \sqrt{x^2+y^2} \geqslant \sqrt{x^2} = |x|$, on obtient

$$|1+z|^2 = (1+z)(1+\bar{z}) = 1 + (z+\bar{z}) + z\bar{z} =$$
$$= 1 + 2x + |z|^2 \leqslant 1 + 2|z| + |z|^2 = (1+|z|)^2,$$

d'où $|1+z| \leqslant 1 + |z|$. Si maintenant $z \neq 0$ et $z' \neq 0$, alors

$$|z+z'| = |z(1+z^{-1}z')| = |z|\cdot|1+z^{-1}z'| \leqslant$$
$$\leqslant |z|\cdot(1+|z^{-1}z'|) = |z|(1+|z|^{-1}|z'|) = |z| + |z'|.$$

Partant des résultats obtenus, nous pouvons énoncer un certain principe général: la forme usuelle (7) des nombres complexes est commode pour exprimer leurs propriétés additives, et la forme trigonométrique (10), pour les propriétés multiplicatives. La non-observation de ce principe conduit à des formules extrêmement compliquées qui voilent le sens des opérations.

4. Elévation à une puissance et extraction de racines.— De la formule (12) pour la multiplication des nombres complexes donnés sous forme trigonométrique, découle une formule appelée *formule de Moivre*

$$[r(\cos\varphi + i \sin\varphi)]^n = r^n (\cos n\varphi + i \sin n\varphi), \qquad (13)$$

valable pour tous les $n \in \mathbb{Z}$ (sous une autre forme: $|z^n| = |z|^n$, $\arg z^n = n \cdot \arg z$). Dans le cas particulier de la formule (13), où $r = 1$, la formule du binôme (1) (voir chap. 1, § 7) et les relations

$$i^2 = -1, \quad i^3 = -i, \quad i^4 = 1, \quad i^{4k+l} = i^l$$

permettent d'obtenir les expressions pour les sinus et les cosinus d'un angle multiple:

$$\begin{aligned} \cos n\varphi &= \sum_{k \geqslant 0} (-1)^k \binom{n}{2k} \cos^{n-2k} \varphi \cdot \sin^{2k} \varphi, \\ \sin n\varphi &= \sum_{k \geqslant 0} (-1)^k \binom{n}{2k+1} \cos^{n-1-2k} \varphi \cdot \sin^{2k+1} \varphi. \end{aligned} \tag{14}$$

A dire vrai, nous avons déjà utilisé le cas particulier de la formule (14) pour $n = 2$ au cours de la démonstration du théorème 2.

REMARQUE. — Soit $e^{\alpha} = \lim\limits_{n \to \infty} \left(1 + \frac{\alpha}{n}\right)^n$. On démontre en Analyse, par décomposition des fonctions de variable complexe en série de puissances, la *formule d'Euler*

$$e^{i\varphi} = \cos \varphi + i \sin \varphi, \tag{15}$$

d'où découlent tous les résultats que nous avons obtenus. Il suffit seulement de remarquer que

$$e^{i\varphi} e^{i\varphi'} = e^{i(\varphi + \varphi')}, \quad (e^{i\varphi})^n = e^{in\varphi}.$$

La forme trigonométrique d'un nombre complexe z se ramène à l'écriture

$$z = |z| \cdot e^{i\varphi}.$$

On voudrait apprendre ensuite à extraire des racines n-ièmes quelconques des nombres complexes, et la question principale qui s'y pose est de savoir si cela est toujours possible. Il s'avère que toujours, et la formule de Moivre fournit au fond une solution complète de ce problème. Soit donné un nombre complexe $z = r (\cos \varphi + i \sin \varphi)$ et nous voulons trouver un nombre $z' = r' (\cos \varphi' + i \sin \varphi')$ tel que $(z')^n = z$. Exprimons $(z')^n$ d'après la formule de Moivre et comparons les modules et les arguments dans les deux membres de l'égalité $(z')^n = z$. Il vient $(r')^n = r$ et $n\varphi' = \varphi + 2\pi k$ (le terme $2\pi k$ résulte du fait que l'argument n'est défini qu'à $2\pi k$ près). Ainsi,

$$r' = \sqrt[n]{r}, \qquad \varphi' = \frac{\varphi + 2\pi k}{n}$$

(par $\sqrt[n]{r}$ on entend la valeur arithmétique de la racine n-ième du nombre réel positif). Donc, la racine $\sqrt[n]{z}$ existe, mais elle n'est pas définie de façon unique. Pour $k = 0, 1, \ldots, n - 1$ on obtient pour z' n valeurs différentes qui représentent toutes les racines pos-

sibles, car il résulte de $k = nq + r$, $0 \leqslant r \leqslant n - 1$, que

$$\varphi' = \frac{\varphi + 2\pi r}{n} + 2\pi q.$$

Nous avons ainsi démontré le théorème suivant:

THÉORÈME 3. — *Il est toujours possible d'extraire la racine n-ième d'un nombre complexe* $z = |z| (\cos \varphi + i \sin \varphi)$. *Toutes les n valeurs de la racine n-ième de z forment les sommets d'un polygone régulier de n côtés inscrit dans un cercle de centre au zéro et de rayon* $\sqrt[n]{|z|}$:

$$\sqrt[n]{z} = \sqrt[n]{|z|} \left(\cos \frac{\varphi + 2\pi k}{n} + i \sin \frac{\varphi + 2\pi k}{n}\right), \tag{16}$$

$k = 0, 1, \ldots, n - 1$. ■

COROLLAIRE. — *Les racines n-ièmes de l'unité s'expriment par la formule*

$$\sqrt[n]{1} = \varepsilon_k = \cos \frac{2\pi k}{n} + i \sin \frac{2\pi k}{n}, \quad k = 0, 1, \ldots, n - 1. \tag{17}$$

Elles forment les sommets d'un polygone régulier de n côtés inscrit dans un cercle de centre au zéro et de rayon 1. ■

On déduit immédiatement de (16) et de (17) que le nombre de racines réelles $\sqrt[n]{z}$ est égal à zéro, à un ou à deux, et celui de racines $\sqrt[n]{1}$, à un ou à deux.

On dit que la racine n-ième de l'unité est *primitive* si elle n'est pas racine m-ième de l'unité, quel que soit $m < n$. L'exemple est fourni par:

$$\varepsilon = \varepsilon_1 = \cos \frac{2\pi}{n} + i \sin \frac{2\pi}{n} \quad \text{et} \quad \varepsilon_{n-1}.$$

Toute autre racine ε_k est une puissance d'une racine primitive:

$$\varepsilon_k = \varepsilon_1^k,$$

ce qui découle toujours de la formule de Moivre. On a de plus, $\varepsilon_k \varepsilon_l = \varepsilon_{k+l}$ si $k + l$ est pris par rapport au module n. En particulier, $\varepsilon_k^{-1} = \varepsilon_{n-k}$, $\varepsilon_0 = 1$. Etant déjà suffisamment calés en théorie des groupes, nous remarquons que *les racines n-ièmes de l'unité engendrent un groupe cyclique* $\langle \varepsilon \rangle$ *d'ordre n*.

Nous avons ainsi obtenu encore une réalisation d'un groupe cyclique d'ordre n. En vertu du théorème 6 (chap. 4, § 3), il y a une bijection entre tous ses sous-groupes et les diviseurs positifs d du nombre n. Pour tout $d \mid n$ il existe dans $\langle \varepsilon \rangle$ un seul sous-groupe $\langle \varepsilon^{\frac{n}{d}} \rangle$ d'ordre d. *La racine* ε_m *est primitive si, et seulement si,* $\langle \varepsilon_m \rangle = \langle \varepsilon \rangle$, *c'est-à-dire si* Card $\langle \varepsilon^m \rangle = n$, *ce qui est possible dans le seul cas où m est premier avec n*. C'est ainsi, par exemple, que pour $n = 12$,

les racines primitives seront ε, ε^5, ε^7, ε^{11}. Dans le cas où $n = p$ est premier, toutes les racines n-ièmes de l'unité, différentes de 1, sont primitives. Du point de vue algébrique, sans tenir compte de la représentation géométrique, toutes les racines primitives n-ièmes sont équivalentes.

En reprenant l'extraction de la racine n-ième d'un nombre complexe arbitraire $z \neq 0$, remarquons que, si z' est une racine fixée quelconque (disons $z' = \sqrt[n]{|z|}\left(\cos\frac{\varphi}{n} + i\sin\frac{\varphi}{n}\right)$), toutes les autres racines sont de la forme $z'\varepsilon_k$, $k = 0, 1, \ldots, n-1$. Cette assertion est en accord avec la formule (16).

5. Théorème d'unicité.— Nous ne pourrons apprécier que par la suite tous les avantages que le corps $\mathbb{C}$ présente par rapport au corps $\mathbb{R}$. Pour le moment, remarquons que le fait de contenir toutes les racines de l'unité justifie, à lui seul, cet intérêt particulier qui est porté aux nombres complexes. Il est naturel de se demander si la famille des corps ayant des propriétés analogues est suffisamment large. Il se trouve qu'on peut énoncer un *théorème sur l'unicité du corps des nombres complexes*, qui est le suivant :

THÉORÈME 4. — *Soit K un corps isomorphe à $\mathbb{R}$ (en particulier, $K = \mathbb{R}$) et soit P une extension de K obtenue par adjonction de la racine j de l'équation $x^2 + 1 = 0$. Alors, P est isomorphe à $\mathbb{C}$.*

DÉMONSTRATION. — Par la définition donnée au chapitre 4, § 4, nº 5, $P = K(j)$ est un sous-corps minimal d'un corps F, qui contient K et j. Le corps F étant donné, nous pouvons considérer les éléments de la forme $a + jb$ avec $a, b \in K$, où le produit et la somme sont entendus au sens des opérations définies dans F. Aux différents couples $a, b \in K$ correspondent des éléments différents $a + jb$, car dans le cas contraire, il existerait un élément égal à zéro $a' + jb'$, avec $a' \neq 0$ ou $b' \neq 0$. Si $b' = 0$, on a aussi évidemment $a' = 0$. Si $b' \neq 0$, on obtient $j = -a'/b' \in K$ ce qui est manifestement absurde: $K \cong \mathbb{R}$, alors que dans $\mathbb{R}$ l'équation $x^2 + 1 = 0$ est insoluble; donc, $j \notin K$. En utilisant la seule égalité $j^2 = -1$ et en opérant dans le corps F, on obtient les formules

$$\begin{aligned}(a_1 + jb_1) + (a_2 + jb_2) &= (a_1 + a_2) + j(b_1 + b_2),\\ (a_1 + jb_1)\cdot(a_2 + jb_2) &= (a_1a_2 - b_1b_2) + j(a_1b_2 + a_2b_1).\end{aligned} \qquad (18)$$

On a, de plus

$$(a + jb)^{-1} = \frac{a}{a^2+b^2} + j\,\frac{-b}{a^2+b^2} \quad \text{pour} \quad a^2+b^2 \neq 0.$$

Cela signifie que l'ensemble $\{a + jb \mid a, b \in K\}$ contenu dans P est stable pour toutes les opérations de F et possède donc une struc-

ture de corps commutatif. P étant minimal, on a l'égalité

$$P = \{a + jb \mid a,\ b \in K\}.$$

En outre, les formules (18) et (9) sont tout à fait identiques.

Si $f\colon K \to \mathbb{R}$ est un isomorphisme donné, l'application

$$f^*\colon a + jb \mapsto (f(a),\ f(b)),$$

qui associe aux éléments du corps P les points du plan $\mathbb{C}$ de coordonnées $f(a)$, $f(b)$, sera, du fait de ce qui précède, un isomorphisme des corps P et $\mathbb{C}$. ■

Au n° 1, nous avons déjà considéré un corps P contenu dans M_2. Or, le nombre de tels corps est évidemment aussi grand que l'on veut (au paragraphe suivant nous considérerons encore un exemple). D'après ce qui a été démontré, tous ces corps sont isomorphes. Remarquons que dans l'énoncé du théorème 4, il aurait fallu écrire $x^2 + \tilde{1} = \tilde{0}$, où $\tilde{1}$ et $\tilde{0}$ sont les éléments unité et zéro du corps K. Par exemple, dans le corps $P \subset M_2$ on a $J^2 + \tilde{1} = \tilde{0}$, où $\tilde{1} = E$ et $\tilde{0}$ est la matrice nulle.

En plus de $\mathbb{Q}$ et $\mathbb{R}$, le corps $\mathbb{C}$ contient beaucoup d'autres sous-corps. Ce sont les extensions du corps $\mathbb{Q}$ obtenues par adjonction d'un élément quelconque de $\mathbb{C}$ non contenu dans $\mathbb{Q}$ qui présentent un intérêt particulier.

EXEMPLE 1 (*corps quadratique*).— Soit d un entier non nul, non nécessairement positif et tel que $\sqrt{d} \notin \mathbb{Q}$. Le corps $\mathbb{Q}(\sqrt{d}) \subset \mathbb{C}$ est dit *quadratique réel* pour $d > 0$, et *quadratique imaginaire* pour $d < 0$. Nous avons fait mention du corps $\mathbb{Q}(\sqrt{2})$ au chapitre 4, § 4. Un raisonnement qui reprend mot par mot la démonstration du théorème 4, où l'on remplace j par $\sqrt{d}$ et la relation $j^2 = -1$ par $(\sqrt{d})^2 = d$, montre que

$$\mathbb{Q}(\sqrt{d}) = \{a + b\sqrt{d} \mid a,\ b \in \mathbb{Q}\}.$$

En particulier, les formules (18) s'écrivent sous la forme

$$\begin{aligned}(a_1 + b_1\sqrt{d}) + (a_2 + b_2\sqrt{d}) &= (a_1 + a_2) + (b_1 + b_2)\sqrt{d},\\ (a_1 + b_1\sqrt{d})(a_2 + b_2\sqrt{d}) &= (a_1a_2 + b_1b_2d) + (a_1b_2 + a_2b_1)\sqrt{d},\end{aligned} \tag{19}$$

et

$$(a + b\sqrt{d})^{-1} = \frac{a}{a^2 - db^2} + \frac{-b}{a^2 - db^2}\sqrt{d}$$

pour $a + b\sqrt{d} \neq 0$ (c'est-à-dire pour a et b simultanément non nuls).

En se servant des formules (19) on vérifie aisément que l'application

$$f\colon a + b\sqrt{d} \mapsto a - b\sqrt{d}$$

est un automorphisme du corps $\mathbb{Q}(\sqrt{d})$ (un analogue de conjugaison complexe). On appelle *norme* d'un nombre $\alpha = a + b\sqrt{d}$ le nombre

$$N(\alpha) = a^2 - db^2 = \alpha f(\alpha).$$

Il est évident que $N(\alpha) = 0 \Leftrightarrow \alpha = 0$. f étant un automorphisme, on a

$$N(\alpha\beta) = \alpha\beta f(\alpha\beta) = \alpha\beta f(\alpha) f(\beta) = \alpha f(\alpha)\cdot\beta f(\beta) = N(\alpha)\cdot N(\beta).$$

En particulier, $N(\alpha)\cdot N(\alpha^{-1}) = N(\alpha\alpha^{-1}) = N(1) = 1$. Voilà pourquoi la norme jouit des propriétés essentielles (du carré) du module dans le corps $\mathbb{C}$.

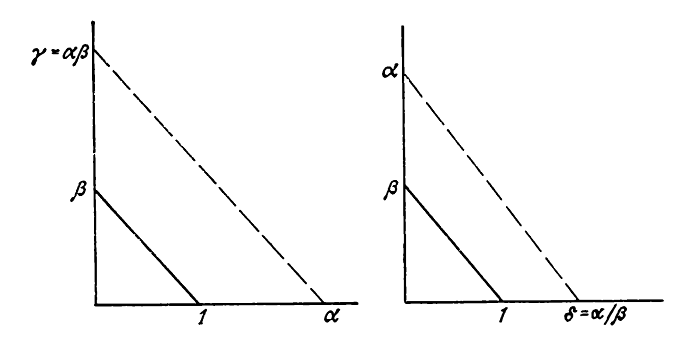
EXEMPLE 2 (*corps des nombres construits*).— Supposons donnés sur le plan cartésien $\mathbb{R}^2$ deux points (0, 0) et (1, 0). Toutes les constructions qui suivent sont réalisées uniquement au moyen d'une règle et d'un compas. Si l'on a construit des points P et Q, on peut évidemment considérer que le segment PQ joignant ces points est aussi construit. Ayant construit un point P et un segment r, on peut également construire un cercle de rayon r et de centre P. Les intersections des droites (segments de droite) et des cercles construits, pris deux à deux, sont également construites.

Un nombre complexe $a + ib \in \mathbb{C}$ est dit *construit* si, partant des points (0, 0) et (1, 0), on peut construire à l'aide d'une suite finie de constructions (admissibles) indiquées plus haut, le point $P = (a, b)$. Il n'est pas difficile d'établir que la propriété du nombre $a + ib$ d'être construit est équivalente à celle de $|a|$ et de $|b|$. L'ensemble des points du plan construits au moyen d'une règle et d'un compas et, par conséquent, l'ensemble de tous les nombres complexes construits sera désigné par le symbole CS.

THÉORÈME 5.— *L'ensemble* CS *est sous-corps du corps* $\mathbb{C}$.

DÉMONSTRATION. — Il résulte immédiatement de la définition d'un nombre construit que l'ensemble CS est stable pour l'opération d'addition et l'opération de passage de $z = a + ib \in$ CS à $-z = -a - ib$.

On construit sur les axes des coordonnées les segments de longueurs 1, α, β et l'on considère les triangles semblables construits sur les figures ci-dessous (qui peuvent admettre de légères modifications)



Il est facile de s'assurer que le produit $\gamma = \alpha\beta$ et le quotient $\delta = \alpha/\beta$ sont aussi dans CS. Puisque la construction de $zz' = (a + ib)(a' + ib') = (aa' - bb') + i(ab' + a'b)$ et de $1/z = a/(a^2 + b^2) + ib/(a^2 + b^2)$ se ramène finalement à celle des grandeurs de type γ et δ, le produit zz' et le quotient $1/z$ appartiennent également à CS. Ceci étant, on a démontré que l'ensemble CS est stable pour toutes les opérations définies dans le corps $\mathbb{C}$.

On convient d'appeler *corps des nombres construits* tout sous-corps $P \subset$ CS. Il est clair que $\mathbb{Q} \subset P$ et que P est un corps de caractéristique nulle.

EXERCICES

1. Trouver tous les nombres complexes z de module 1, tels que $z^2 + (1+i)\,z$ prenne des valeurs imaginaires pures. Représenter ces nombres par des points du plan $\mathbb{C}$.

2. Que peut-on dire du corps $\mathbb{R}(\delta)$ obtenu de $\mathbb{R}$ par adjonction d'un nombre complexe δ qui satisfait à l'égalité $\delta^4 = -1$?

3. Soient donnés A, $B \in M_n(\mathbb{R})$. Démontrer en se servant du théorème 1 que $\det \overline{(A + iB)} = \det (A - iB)$ (le trait désigne un nombre conjugué).

4. Soient A, $B \in M_n(\mathbb{R})$,

$$C = \left\| \begin{matrix} A & -B \\ B & A \end{matrix} \right\| \in M_{2n}(\mathbb{R}).$$

En appliquant à la matrice réelle C les transformations élémentaires de types $(\overline{\mathrm{I}})$ et $(\overline{\mathrm{II}})$ sur le corps des nombres complexes $\mathbb{C}$, montrer que

$$\det C = |\det (A + iB)|^2.$$

5. (G. Polya et G. Szego.) En utilisant les exercices 3 et 4, expliquer un fait « étrange » suivant. Un système linéaire carré homogène

$$\begin{aligned} d_{11}z_1 + \ldots + d_{1n}z_n &= 0, \\ \ldots\ldots\ldots\ldots \\ d_{n1}z_1 + \ldots + d_{nn}z_n &= 0 \end{aligned} \qquad (*)$$

à coefficients complexes $d_{kl} = a_{kl} + ib_{kl}$ et à inconnues $z_l = x_l + iy_l$ possède une solution non triviale $(z_1, \ldots, z_n)$ si, et seulement si, $\det (d_{kl}) = a + ib = 0$ (voir les remarques générales faites à ce propos au chap. 4, § 4, nº 7). Cette condition entraîne deux équations $a = 0$, $b = 0$ qui lient $2n^2$ grandeurs réelles a_{kl}, b_{kl}. D'autre part, le système $(*)$ peut être présenté sous forme d'un système de $2n$ équations linéaires homogènes à $2n$ inconnues réelles x_l, y_l. Maintenant, la condition d'existence de solution non triviale s'écrit sous la forme de l'égalité à zéro d'un déterminant réel d'ordre $2n$, ce qui donne une seule équation liant a_{kl}, b_{kl}. Comment peut-on mettre en accord ces deux résultats?

6. Ayant en vue que les automorphismes du corps quadratique $\mathbb{Q}(\sqrt{d})$ doivent laisser invariants les nombres rationnels, trouver les automorphismes de ce corps. (Réponse. L'application identique et l'application $a + b\sqrt{d} \mapsto a - b\sqrt{d}$.)

7. Quelle est la somme de toutes les racines n-ièmes de l'unité avec $n > 1$? Que peut-on dire de la somme des racines primitives 12-ième et 15-ième?

8. Trouver et représenter géométriquement le noyau et l'image de l'application $(\mathbb{R}, +) \to (\mathbb{C}^*, \cdot)$, $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$ définie par $t \mapsto e^{2\pi i t}$ (voir formule (15)).

9. Montrer que $\xi = \dfrac{2+i}{2-i}$ n'est pas racine de 1, bien que $|\xi| = 1$. (Indication. $\xi^n = 1 \Rightarrow (2 - i)^n = (2 + i)^n = (2 - i + 2i)^n = (2 - i)^n + \ldots + (2i)^n \Rightarrow (2 - i)(a + ib) = (2i)^n \Rightarrow 5 \cdot (a^2 + b^2) = 2^{2n} \Rightarrow 5 \mid 2^{2n}$, contradiction.)

10. Montrer que

$$\begin{vmatrix} x_0 & x_1 & \ldots & x_{n-1} \\ x_1 & x_2 & \ldots & x_0 \\ \ldots & \ldots & \ldots & \ldots \\ x_{n-1} & x_0 & \ldots & x_{n-2} \end{vmatrix} = \prod_{i=0}^{n-1} (x_0 + \zeta^i x_1 + \zeta^{2i} x_2 + \ldots + \zeta^{(n-1)i} x_{n-1}),$$

où ζ est la racine primitive n-ième de 1.

§ 2. Anneau des polynômes

Les polynômes constituent, de même que les systèmes linéaires que nous avons considérés aux chapitres 2 et 3, un domaine bien étudié de l'algèbre traditionnelle. De nombreux problèmes mathématiques, les plus variés, sont énoncés et résolus en termes de polynômes. Ceci s'explique par de nombreuses causes dont l'une est la propriété d'universalité dont jouit l'anneau des polynômes et qui sera sommairement étudiée aux n°s 1 et 2 du présent paragraphe.

Soient K un anneau unitaire commutatif (et, comme à l'ordinaire, associatif) et A un sous-anneau contenant 1. Si $t \in K$, le sous-anneau minimal de K, contenant A et t, sera manifestement constitué par les éléments de la forme

$$a(t) = a_0 + a_1t + a_2t^2 + \ldots + a_nt^n, \qquad (*)$$

où $a_s \in A$, $n \in \mathbb{Z}$, $n \geqslant 0$. Nous le désignons par le symbole $A[t]$ et appellerons anneau obtenu à partir de A par adjonction de l'élément t. Quant à l'expression $(*)$, on appellera polynôme en t à coefficients dans A. Des exemples bien simples

$$a(t) + b(t) = (a_0 + a_1t + a_2t^2) + (b_0 + b_1t + b_2t^2) =$$
$$= (a_0 + b_0) + (a_1 + b_1)\,t + (a_2 + b_2)\,t^2,$$
$$a(t)\cdot b(t) = a_0b_0 + (a_0b_1 + a_1b_0)\,t +$$
$$+ (a_0b_2 + a_1b_1 + a_2b_0)\,t^2 + (a_1b_2 + a_2b_1)\,t^3 + a_2b_2t^4$$

montrent ce qu'on doit entendre par la somme et le produit des polynômes. Il est évident que la réduction des termes semblables est basée sur le fait que tous les éléments a_i, b_j, t^k sont deux à deux permutables.

Maintenant il est temps de se rappeler que t est un élément, pris au hasard, de l'anneau K. Voilà pourquoi, les expressions $(*)$, ayant une forme différente, peuvent en fait s'identifier. Si, par exemple, $A = \mathbb{Q}$, $t = \sqrt{2}$, alors $t^2 = 2$ et $t^3 = 2t$ sont des relations qui ne découlent aucunement des règles formelles. Pour venir à la notion habituelle de polynôme, il faut s'affranchir de toutes les relations incidentes de ce genre. A cet effet, il convient d'entendre par t un symbole arbitraire non nécessairement contenu dans K. Ce dernier est appelé à jouer un rôle purement auxiliaire. Ce sont les règles par lesquelles on obtient les coefficients des expressions $a(t) + b(t)$, $a(t)\,b(t)$ qui revêtent une importance beaucoup plus grande. Ayant en vue ces remarques préliminaires, passons à la définition rigoureuse d'un être algébrique appelé polynôme et à l'ensemble de tels êtres, c'est-à-dire à l'anneau des polynômes.

1. Polynômes à une indéterminée.— Soit A un anneau unitaire commutatif arbitraire. On forme un nouvel anneau B dont les élé-

ments sont des suites ordonnées infinies

$$f = (f_0, f_1, f_2, \dots), \quad f_i \in A, \tag{1}$$

telles que tous les f_i, sauf un nombre fini de ses éléments, sont nuls. Définissons dans l'ensemble B les opérations d'addition et de multiplication en posant

$$f + g = (f_0, f_1, f_2, \dots) + (g_0, g_1, g_2, \dots) =$$
$$= (f_0 + g_0, f_1 + g_1, f_2 + g_2, \dots),$$
$$f \cdot g = h = (h_0, h_1, h_2, \dots),$$

où

$$h_k = \sum_{i+j=k} f_i g_j, \quad k = 0, 1, 2, \dots$$

Il est clair qu'à la suite de l'addition et de la multiplication on obtient de nouveau des suites de la forme (1) contenant un nombre fini de termes non nuls, c'est-à-dire des éléments de B. La vérification de tous les axiomes de l'anneau (voir chap. 4, § 4), excepté peut-être l'axiome d'associativité, est évidente. En effet, puisque l'addition de deux éléments de B se ramène à l'addition d'un nombre fini d'éléments de l'anneau A, $(B, +)$ est un groupe commutatif qui admet un élément zéro $(0, 0, 0, \dots)$ et un élément $-f = (-f_0, -f_1, -f_2, \dots)$ opposé à un élément quelconque $f = (f_0, f_1, f_2, \dots)$. La commutativité de la multiplication résulte directement de la symétrie de l'expression des éléments h_k par f_i et g_j. Cette expression montre aussi que les éléments de B vérifient la loi de distributivité $(f + g)h = fh + gh$. Quant à l'associativité de l'opération de multiplication, considérons trois éléments quelconques

$$f = (f_0, f_1, f_2, \dots), \quad g = (g_0, g_1, g_2, \dots), \quad h = (h_0, h_1, h_2, \dots)$$

de l'ensemble B. On a $fg = d = (d_0, d_1, d_2, \dots)$, où $d_l = \sum_{i+j=l} f_i g_j$, $l = 0, 1, 2, \dots$, et $(fg)h = dh = e = (e_0, e_1, e_2, \dots)$, où $e_s = \sum_{l+k=s} d_l h_k = \sum_{l+k=s} (\sum_{i+j=l} f_i g_j) h_k = \sum_{i+j+k=s} f_i g_j h_k$. Le calcul de $f(gh)$ donne le même résultat. Ainsi, B est un *anneau commutatif et associatif à élément unité* $(1, 0, 0, \dots)$.

On additionne et multiplie les suites $(a, 0, 0, \dots)$ de la même façon que les éléments de l'anneau A. Cela permet d'identifier de telles suites avec les éléments correspondants de A, c'est-à-dire poser $a = (a, 0, 0, \dots)$ pour tous les $a \in A$. A devient par là même un sous-anneau de l'anneau B. Désignons $(0, 1, 0, 0, \dots)$ par X et appelons X *indéterminée* sur A. En utilisant l'opération de mul-

tiplication introduite dans B, on obtient

$$\begin{aligned} X &= (0, 1, 0, 0, \ldots), \\ X^2 &= (0, 0, 1, 0, \ldots), \qquad (2) \\ &\ldots\ldots\ldots\ldots\ldots \\ X^n &= (0, 0, \ldots, 0, 1, 0, \ldots). \end{aligned}$$

En outre, compte tenu de (2) et de l'inclusion $A \subset B$, on a

$$(0, 0, \ldots, 0, a, 0, \ldots) = aX^n = X^n a.$$

Ainsi donc, si f_n est le dernier terme non nul de la suite $f = (f_0, f_1, \ldots, f_n, 0, 0, \ldots)$, on peut écrire avec les nouvelles notations

$$\begin{aligned} f &= (f_0, \ldots, f_{n-1}, 0, 0, \ldots) + f_n X^n = \\ &= (f_0, \ldots, f_{n-2}, 0, 0, \ldots) + f_{n-1}X^{n-1} + f_n X^n = \\ &= f_0 + f_1 X + f_2 X^2 + \ldots + f_n X^n. \qquad (3) \end{aligned}$$

Une telle représentation de l'élément f est univoque parce que $f_0, \ldots, f_n$ figurant dans le second membre de l'égalité (3) sont les termes de la suite $(f_0, \ldots, f_n, 0, \ldots)$ qui est nulle si, et seulement si, $f_0 = \ldots = f_n = 0$.

DÉFINITION. — *L'anneau B introduit plus haut se note $A[X]$ et s'appelle anneau des polynômes à une indéterminée X sur A ; ses éléments s'appellent polynômes.*

Le nom d'indéterminée ou de variable donné à une lettre fixe X ne peut certes être considéré comme une invention terminologique bien heureuse mais, ne provoquant aucun malentendu, cette appellation a été consacrée par l'usage.

C'est à dessein de ne pas confondre notre polynôme $f = X$ spécialement formé et une variable x utilisée en théorie des fonctions et qui parcourt un certain ensemble de valeurs, que nous avons introduit la lettre X majuscule (convention purement temporaire qui ne sera pas obligatoire par la suite). Une écriture plus habituelle du polynôme f se présente sous la forme

$$f(X) = a_0 X^n + a_1 X^{n-1} + \ldots + a_n,$$

c'est-à-dire suivant les puissances décroissantes de X. Dans la suite de cet ouvrage, nous utiliserons celle des écritures qui semble être la plus commode. Les éléments f_i (et a_i) s'appellent *coefficients* du polynôme f. Le polynôme f est dit *nul* si tous ses coefficients sont égaux à zéro. Le coefficient f_0 de X à la puissance zéro est encore appelé *terme constant*. Si $f_n \neq 0$, f_n s'appelle coefficient *dominant* et n, *degré* du polynôme et l'on écrit $n = \deg f$. Au polynôme nul on

attribue un degré symbolique noté $-\infty$ $(-\infty + (-\infty) = -\infty$, $-\infty + n = -\infty$, $-\infty < n$ pour tout $n \in \mathbb{N})$. Les polynômes de degrés 1, 2, 3, . . . sont dits respectivement *linéaires*, *quadratiques* (ou *carrés*), *cubiques*, etc.

Le rôle de l'unité de l'anneau $A\,[X]$ est joué par l'élément unité 1 de l'anneau A, considéré comme un polynôme de degré zéro. Il s'ensuit de la définition même des opérations d'addition et de multiplication dans $A\,[X]$ que

$$\deg (f + g) \leqslant \max (\deg f, \deg g), \quad \deg (fg) \leqslant \deg f + \deg g, \tag{4}$$

quels que soient les polynômes

$$f = f_0 + f_1X + \ldots + f_nX^n, \quad g = g_0 + g_1X + \ldots + g_mX^m \tag{5}$$

de degrés respectifs n et m.

La deuxième des inégalités (4) est en fait remplacée par l'égalité

$$\deg (fg) = \deg f + \deg g$$

chaque fois que le produit f_ng_m des coefficients dominants des polynômes (5) est différent de zéro. En effet

$$fg = f_0g_0 + (f_0g_1 + f_1g_0)\,X + \ldots + (f_ng_m)\,X^{n+m}. \tag{6}$$

Or, cela signifie qu'on peut énoncer le théorème suivant:

THÉORÈME 1. — *Si A est un anneau intègre, $A\,[X]$ est, lui aussi, un anneau intègre.* ■

La place que l'anneau des polynômes occupe parmi les anneaux commutatifs, est partiellement illustrée par le théorème suivant:

THÉORÈME 2. — *Soit K un anneau commutatif contenant A comme sous-anneau. Alors, pour tout élément $t \in K$, il existe un seul homomorphisme d'anneaux $\Pi_t : A\,[X] \to K$ tel que*

$$\Pi_t (a) = a, \ \forall a \in A, \ \Pi_t (X) = t. \tag{7}$$

DÉMONSTRATION. — Supposons d'abord qu'un tel homomorphisme Π_t existe. Puisque $\Pi_t (f_i) = f_i$ pour tout coefficient du polynôme f écrit sous la forme standard (3), et $\Pi_t (X^k) = (\Pi_t (X))^k = t^k$ (propriété de l'homomorphisme et condition (7)), il vient

$$\Pi_t (f) = \Pi_t (f_0 + f_1X + \ldots + f_nX^n) =$$
$$= f_0 + f_1t + \ldots + f_nt^n, \tag{8}$$

ce qui signifie que $\Pi_t (f)$ est déterminé univoquement et s'exprime par la formule (8). Inversement, en définissant l'application Π_t par la formule (8), on satisfait manifestement à la condition (7) et l'on obtient l'homomorphisme des anneaux. Cette dernière proposition est évidente pour l'application Π_t des groupes additifs des anneaux. En ce qui concerne la multiplication, l'application Π_t

opérant sur le produit (6), l'utilisation de la loi (générale) de distributivité donne

$$\Pi_t(fg) = f_0g_0 + (f_0g_1 + f_1g_0)\,t + \ldots + (f_ng_m)\,t^{n+m} =$$
$$= \Big(\sum_{i=0}^{n} f_it^i\Big)\Big(\sum_{j=0}^{m} g_jt^j\Big) = \Pi_t(f)\cdot\Pi_t(g). \ \blacksquare$$

L'image d'un polynôme $f = f(X)$ par l'application Π_t, définie par (8), s'appelle substitution de t à X dans f ou (ce qui est moins rigoureux) tout simplement valeur de f pour $X = t$, de sorte que $\Pi_t(f) = f(t)$. Connaître $\Pi_t(f)$ c'est savoir calculer la valeur de f pour $X = t$. Les homomorphismes Π_x, $x \in A$, servent de chaînon entre les points de vue fonctionnel et algébrique sur le polynôme. Par définition, un polynôme linéaire $X - c = (-c, 1, 0, \ldots)$ n'est jamais nul, alors que la fonction $x \mapsto x - c$ qui lui est associée, s'annule pour $x = c$. Un autre exemple: un polynôme non nul $X^2 + X$ à coefficients dans le corps $\mathbb{F}_2$ (où $1 + 1 = 0$) correspond à une fonction nulle $\tilde{f}\colon \mathbb{F}_2 \to \mathbb{F}_2$, car $0^2 + 0 = 0$ et $1^2 + 1 = 0$.

On dit qu'un élément $t \in K$ est *algébrique sur* A si $\Pi_t(f) = 0$ pour un certain $f \in A[X]$. Si $\Pi_t\colon A[X] \to K$ est un plongement isomorphe (un monomorphisme), t est un élément *transcendant sur* A. Dans le cas où $A = \mathbb{Q}$ et $K = \mathbb{C}$, on parle tout simplement des *nombres algébriques* et *transcendants*. Comme exemples de nombres transcendants on peut indiquer e et π définis en Analyse, alors que les exemples de nombres algébriques sont $\sqrt{2}$, $\sqrt{3}$, $\sqrt{2} + \sqrt{3}$.

Pour mesurer l'écart que l'anneau $A[t] \subset K$ obtenu au début de ce paragraphe présente par rapport à l'anneau des polynômes $A[X]$, considérons le noyau $J_t = \operatorname{Ker} \Pi_t$ de l'homomorphisme Π_t du théorème 2. Suivant (7), Π_t est une transformation identique de A et donc $A \cap J_t = 0$. Remarquons, en passant, que $J_t = 0$ si t est un élément transcendant sur A. D'après le théorème relatif aux homomorphismes d'anneaux (chap. 4, § 4, n° 4, théorème 2) on a

$$A[t] \cong A[X]/J_t. \tag{9}$$

L'isomorphisme (9) sert au fond à exprimer la *propriété universelle* de l'anneau des polynômes $A[X]$. Elle sera mieux mise en évidence par l'assertion suivante qui généralise le théorème 2.

THÉORÈME 3. — *Soient A et K deux anneaux commutatifs quelconques, t un élément de K et $\varphi\colon A \to K$ un homomorphisme. Il existe alors un prolongement et un seul de φ à l'homomorphisme $\varphi_t\colon A[X] \to K$ de l'anneau des polynômes $A[X]$ dans K, qui transforme l'indéterminée X en t.*

La démonstration de ce théorème n'étant que légèrement différente de celle du théorème 2, nous laissons au lecteur le soin de la faire à titre d'exercice. $\blacksquare$

2. Polynômes à plusieurs indéterminées.— Si, dans la situation $A \subset K$ considérée au début de ce paragraphe, on prend n éléments arbitraires $t_1, \ldots, t_n \in K$ et on considère dans K l'intersection de tous les sous-anneaux contenant $A, t_1, \ldots, t_n$, on obtient un anneau $A[t_1, \ldots, t_n]$. L'écriture formelle de ses éléments suggère, comme dans le cas où $n = 1$, qu'il est nécessaire d'introduire un anneau des polynômes à n indéterminées. Cela se fait de façon très simple. Rappelons que la construction de l'anneau $B = A[X]$ contenait un anneau unitaire commutatif arbitraire A. Nous pouvons maintenant remplacer dans notre construction l'anneau A par B et construire un anneau $C = B[Y]$, où Y est une nouvelle indéterminée indépendante qui joue par rapport à B le même rôle que celui de X par rapport à A. Les éléments de C s'écrivent d'une manière et d'une seule sous la forme $\sum b_j Y^j$, $b_j \in B$, et B s'identifie avec un sous-anneau de C, à savoir avec l'ensemble des éléments $bY^0 = b \cdot 1$. Puisque $b_j = \sum a_{ij} X^i$ est également une écriture unique des éléments $b_j \in B$, tout élément de C est de la forme

$$\sum_{i=0}^{k} \sum_{j=0}^{l} a_{ij} X^i Y^j, \quad a_{ij} \in A,$$

dans laquelle on sous-entend (d'après le sens de la construction) que les a_{ij} sont permutables avec X et Y, alors que l'indéterminée X est permutable avec Y. L'anneau C s'appelle anneau des polynômes à deux indéterminées (à deux variables) X et Y sur A.

En répétant cette construction un nombre de fois suffisant, on obtient un *anneau* $A[X_1, \ldots, X_n]$ *des polynômes à* n *indéterminées* (ou *variables*) $X_1, \ldots, X_n$ *sur* A.

Convenons de noter par (i) une suite $(i_1, \ldots, i_n) \in \overline{\mathbb{N}}^n$ de n entiers non négatifs $i_1, \ldots, i_n$ ($\overline{\mathbb{N}} = \mathbb{N} \cup \{0\}$). Alors, tout élément $f \in A[X_1, \ldots, X_n]$ s'écrit sous la forme

$$f = \sum_{(i)} a_{(i)} X^{(i)}, \quad a_{(i)} \in A, \tag{10}$$

où $X^{(i)} = X_1^{i_1} \ldots X_n^{i_n}$ est un *monôme*, si bien que f est une combinaison linéaire de monômes à coefficients dans A. En vertu de la définition des polynômes, tous les coefficients $a_{(i)}$ figurant dans (10) sont nuls à l'exception d'un nombre fini d'entre eux. L'unicité de l'écriture (10) découle directement de l'assertion suivante:

Un polynôme f *est nul si, et seulement si, tous ses coefficients* $a_{i_1 \ldots i_n}$ *sont nuls.* Pour $n = 1$, cela a été déjà constaté lors de la construction de l'anneau $A[X]$. Pour $n > 1$, on peut le plus simplement utiliser un raisonnement par récurrence sur n. A savoir, nous pouvons écrire

$$f = \sum a_{i_1 \ldots i_n} X_1^{i_1} \ldots X_n^{i_n} = \sum_{i_n} b_{i_n} X_n^{i_n},$$

où

$$b_{i_n} = \sum_{i_1, \ldots, i_{n-1}} a_{i_1 \ldots i_{n-1} i_n} X_1^{i_1} \ldots X_{n-1}^{i_{n-1}}$$

sont des polynômes à un plus petit nombre d'indéterminées. L'assertion pour $n = 1$ et l'hypothèse de récurrence montrent que

$$f = 0 \Leftrightarrow b_{i_n} = 0, \quad \forall i_n \Leftrightarrow a_{i_1 \ldots i_{n-1} i_n} = 0, \quad \forall (i_1, \ldots, i_n).$$

Maintenant, il est naturel de poser que deux polynômes $f, g \in A[X_1, \ldots, X_n]$ sont *égaux* ou *identiques* si leurs coefficients de monômes égaux coïncident (du fait de ce qui précède : $(i_1, \ldots, i_n) \neq (j_1, \ldots, j_n) \Rightarrow X_1^{i_1} \ldots X_n^{i_n} \neq X_1^{j_1} \ldots X_n^{j_n}$).

On appelle *degré du polynôme f par rapport à l'indéterminée* X_k et l'on note $\deg_k f$ le plus grand des entiers qui se rencontrent parmi les exposants de X_k dans $a_{(i)}X^{(i)}$, avec $a_{(i)} \neq 0$. C'est ainsi, par exemple, que le polynôme $1 + X + XY^3 + X^2Y^2$ est de degré 2 par rapport à X et de degré 3 par rapport à Y. On appelle *degré total* (ou plus simplement *degré*) *d'un monôme* $X_1^{i_1} \ldots X_n^{i_n}$ l'entier $i_1 + \ldots + i_n$. Le *degré total* (ou *degré* tout court) $\deg f$ *d'un polynôme f* sera le plus grand des degrés de ses monômes. On convient que $\deg 0 = -\infty$. Dans le cas des polynômes, la notion de terme dominant suivant le degré n'a pas de sens parce que de tels termes (monômes) peuvent être plusieurs.

De nombreux résultats obtenus au n° 1 pour l'anneau $A[X]$ peuvent être étendus à l'anneau $A[X_1, \ldots, X_n]$. C'est ainsi, par exemple, qu'en partant du théorème 1 et en utilisant le raisonnement par récurrence sur n, on s'assure tout de suite que le théorème suivant est vrai :

THÉORÈME 1'. — *Si A est un anneau intègre, l'anneau* $A[X_1, \ldots, X_n]$ *est, lui aussi, intègre. En particulier, l'anneau des polynômes à n indéterminées sur tout corps P est intègre.* ∎

Soient A un sous-anneau d'un anneau commutatif K et $t_1, \ldots, t_n$ les éléments de K. Alors, la correspondance

$$\Pi_{t_1, \ldots, t_n} : f(X_1, \ldots, X_n) \mapsto f(t_1, \ldots, t_n), \quad \forall f \in A[X_1, \ldots, X_n],$$

définit un homomorphisme $A[X_1, \ldots, X_n] \to K$ (comparer avec le théorème 2). On dit alors que l'on a affaire à une *substitution* de $t_1, \ldots, t_n$ dans f ou qu'il s'agit de la valeur de f pour $X_1 = t_1, \ldots, X_n = t_n$. Si $\operatorname{Ker} \Pi_{t_1, \ldots, t_n} = 0$, on dit que *les éléments* $t_1, \ldots, t_n$ *de l'anneau K sont algébriquement indépendants sur A*. Dans le cas où les éléments $t_1, \ldots, t_n$ sont algébriquement dépendants, il existe un polynôme $f \in A[X_1, \ldots, X_n]$ non nul pour lequel $f(t_1, \ldots, t_n) = 0$.

Enfin, au théorème 3 correspond le théorème analogue suivant :

THÉORÈME 3′ (universalité de l'anneau des polynômes). — *Soient A et K deux anneaux commutatifs, $t_1, \ldots, t_n$ les éléments de K et $\varphi: A \to K$ un homomorphisme des anneaux. Alors, il existe un prolongement et un seul de φ à l'homomorphisme $\varphi_{t_1, \ldots, t_n}: A[X_1, \ldots, X_n] \to K$ qui transforme X_i en t_i, $1 \leqslant i \leqslant n$.*

DÉMONSTRATION. — On démontre le théorème parallèlement à la construction de l'anneau $A[X_1, \ldots, X_n]$, c'est-à-dire par récurrence. En partant du théorème 3, il est naturel de supposer que nous avons un homomorphisme $\varphi_{t_1, \ldots, t_{n-1}}: A[X_1, \ldots, X_{n-1}] \to K$ qui prolonge φ et tel que $\varphi_{t_1, \ldots, t_{n-1}}(X_i) = t_i$, $1 \leqslant i \leqslant n-1$. En remplaçant dans le théorème 3 l'anneau A par l'anneau $A[X_1, \ldots, X_{n-1}]$ et l'homomorphisme φ par $\varphi_{t_1, \ldots, t_{n-1}}$ ainsi qu'en utilisant le fait que $A[X_1, \ldots, X_n] = A[X_1, \ldots, X_{n-1}] \times \times [X_n]$, nous trouvons l'homomorphisme cherché $\varphi_{t_1, \ldots, t_n} = = (\varphi_{t_1, \ldots, t_{n-1}})_{t_n}$ qui transforme X_n en t_n. L'unicité de $\varphi_{t_1, \ldots, t_n}$ n'exige aucune vérification parce que $\varphi_{t_1, \ldots, t_n}$ est entièrement défini par l'opération dans A et sur les éléments $X_1, \ldots, X_n$ qui engendrent $A[X_1, \ldots, X_n]$. ∎

COROLLAIRE. — *A toute permutation $\pi \in S_n$ opérant sur l'ensemble $\{1, 2, \ldots, n\}$ correspond un seul automorphisme $\tilde{\pi}: f \mapsto \pi f$ de l'anneau $A[X_1, \ldots, X_n]$, identique sur A et tel que*

$$(\pi f)(X_1, \ldots, X_n) = f(X_{\pi^{-1}(1)}, \ldots, X_{\pi^{-1}(n)}).$$

DÉMONSTRATION. — Dans l'énoncé du théorème 3′, posons $K = = A[X_1, \ldots, X_n]$, $t_1 = X_{\pi^{-1}(1)}, \ldots, t_n = X_{\pi^{-1}(n)}$ et prenons pour φ la restriction de l'application identique e_K à A. Il en résulte un homomorphisme $\tilde{\pi} = \varphi_{t_1, \ldots, t_n}$ de l'anneau $A[X_1, \ldots, X_n]$ dans lui-même (c'est-à-dire un endomorphisme), qui est de toute évidence un automorphisme. ∎

Une précision utile du théorème 1′ est fournie par le théorème suivant.

THÉORÈME 4. — *Soient f et g deux polynômes quelconques à n indéterminées sur un anneau intègre A.*

Alors on a

$$\deg(fg) = \deg f + \deg g.$$

DÉMONSTRATION. — Appelons *polynôme homogène* ou *forme* de degré m un polynôme $h(X_1, \ldots, X_n)$ dont tous les termes sont de même degré total m. Les formes de degrés 1, 2, 3 s'appellent respectivement formes *linéaires*, *quadratiques* et *cubiques*. Groupant dans f tous les monômes de même degré à coefficients non nuls, nous pouvons représenter univoquement le polynôme $f = \sum a_{(i)} X^{(i)}$ sous la forme de la somme de plusieurs formes f_m de différents degrés :

$$f = f_0 + f_1 + \ldots + f_k, \quad k = \deg f.$$

Si, maintenant,

$$g = g_0 + g_1 + \ldots + g_l, \quad l = \deg g,$$

il est évident que

$$fg = f_0 g_0 + (f_0 g_1 + f_1 g_0) + \ldots + f_k g_l$$

(ce qui ressemble à la relation (6), mais f_i et g_j y ont un autre sens), d'où $\deg fg \leqslant k + l$. D'après le théorème 1', $f_k \neq 0$, $g_l \neq 0$ entraînent $f_k g_l \neq 0$, c'est-à-dire $\deg (fg) = \deg (f_k g_l) = k + l = \deg f + \deg g$. ■

3. Division euclidienne des polynômes.— Les anneaux des polynômes à une indéterminée d'une part et les anneaux des polynômes à plusieurs indéterminées d'autre part présentent non seulement des propriétés communes que nous avons minutieusement soulignées au nº 2, mais aussi des différences essentielles. Si l'on se reporte à la description des idéaux d'un anneau des polynômes, on peut tout de suite mettre en évidence l'une de ces différences. Nous avons vu (chap. 4, § 3, nº 3) que dans l'anneau $\mathbb{Z}$ tout idéal est principal, c'est-à-dire est de la forme $m\mathbb{Z}$. La démonstration de ce fait se basait sur la comparaison des nombres d'après leur valeur à l'aide d'un mécanisme appelé *algorithme de division euclidienne*, décrit pour $\mathbb{Z}$ encore au chapitre 1, § 8, nº 3. Il se trouve qu'un algorithme tout à fait analogue peut être appliqué à l'anneau $A[X]$ sur un anneau intègre A (pour $A = \mathbb{R}$, on le connaît en fait du cours d'algèbre élémentaire: rappelez-vous la division des polynômes!).

THÉORÈME 5. — *Soient A un anneau intègre et g un polynôme dans $A[X]$ dont le coefficient dominant est inversible dans A. Alors, à chaque polynôme $f \in A[X]$ est associé un couple et un seul de polynômes q, $r \in A[X]$ tels que*

$$f = qg + r, \quad \deg r < \deg g. \tag{11}$$

DÉMONSTRATION. — Soit

$$f = a_0 X^n + a_1 X^{n-1} + \ldots + a_n,$$
$$g = b_0 X^m + b_1 X^{m-1} + \ldots + b_m,$$

où $a_0 b_0 \neq 0$ et $b_0 \mid 1$. Raisonnons par récurrence sur n. Si $n = 0$ et $m = \deg g > \deg f = 0$, posons $q = 0$, $r = f$, et $r = 0$, $q = a_0 b_0^{-1}$ si $n = m = 0$. Supposons le théorème démontré pour tous les polynômes de degré $<n$ $(n > 0)$. Sans restreindre la généralité, admettons que $m \leqslant n$; dans le cas contraire prenons $q = 0$ et $r = f$. S'il en est ainsi, on a

$$f = a_0 b_0^{-1} X^{n-m} \cdot g + \bar{f},$$

où $\deg \bar{f} < n$. En raisonnant par récurrence nous pouvons aussi trouver $\bar{q}$ et r tels que $\bar{f} = \bar{q}g + r$ et $\deg r < m$. En posant

$$q = a_0 b_0^{-1} X^{n-m} + \bar{q},$$

nous trouvons un couple de polynômes présentant les propriétés requises.

En nous reportant à la propriété d'unicité du *quotient* q et du *reste* r, supposons que

$$qg + r = f = q'g + r'.$$

Il vient alors $(q' - q)\, g = r - r'$. Suivant le théorème 1, on a $\deg (r - r') = \deg (q' - q) + \deg g$, ce qui ne peut avoir lieu, dans nos conditions, que si $r' = r$ et $q' = q$ (rappelons que $\deg 0 = -\infty$ et que $-\infty + m = -\infty$).

Les raisonnements qui précèdent montrent que les coefficients du quotient q et du reste r appartiennent au même anneau intègre A, c'est-à-dire $f,\ g \in A[X] \Rightarrow q,\ r \in A[X]$. ■

REMARQUE. — Le processus de division euclidienne du polynôme f par le polynôme g se trouve simplifié si g est un *polynôme unitaire*, c'est-à-dire si son coefficient dominant est égal à l'unité. La divisibilité du polynôme f par le polynôme unitaire g est équivalente à la condition que le reste r de la division euclidienne de f par g est nul.

COROLLAIRE. — *Tous les idéaux d'un anneau des polynômes $P[X]$ sur un corps P sont principaux.*

DÉMONSTRATION. — Soit T un idéal non nul de $P[X]$. Choisissons dans T un polynôme $t = t(X)$ de degré minimal. Si f est un polynôme quelconque de T, la division euclidienne de f par t (P étant un corps, on n'a pas besoin de se soucier que le coefficient dominant de $t(X)$ soit inversible) donne l'égalité $f = qt + r$, $\deg r < \deg t$. Il en résulte que $r \in T$, puisque f, t, qt sont éléments de l'idéal. Du fait du choix de t, il ne nous reste qu'à conclure que $r = 0$. Cela signifie que $f(X)$ est divisible par $t(X)$ et que $T = (t) = tP[X]$, ce qui veut dire que T est formé des polynômes divisibles par $t(X)$. ■

Quant aux anneaux des polynômes à plusieurs indéterminées, on peut affirmer *a priori* que même dans $\mathbb{R}[X, Y]$ il existe des idéaux qui ne sont pas principaux.

EXEMPLE. — L'ensemble

$$T = \{Xf + Yg \mid f,\ g \in \mathbb{R}[X, Y]\},$$

formé de polynômes $h(X, Y)$ tels que $h(0, 0) = 0$, est évidemment un idéal dans $\mathbb{R}[X, Y]$. Puisque $1 \in \mathbb{R}[X, Y]$, $T = t(X, Y)\,\mathbb{R}[X, Y]$ entraînerait $t(X, Y) \in T$. C'est pourquoi $t(0, 0) = 0$ et donc $\deg t \geqslant 1$. En appliquant

maintenant le théorème 4 aux égalités

$$X = tu, \qquad Y = tv,$$

on trouve que $\deg u = \deg v = 0$, c'est-à-dire $u, v \in \mathbb{R}$ et $Y = u^{-1}vX$: contradiction qui démontre que l'idéal T n'est pas principal.

Le corollaire du théorème 5 s'avère commode pour une description explicite de l'isomorphisme (9). A titre d'exemple, démontrons une assertion qui complète le théorème 4 du § 1.

THÉORÈME 6. — *Le corps* $\mathbb{C}$ *des nombres complexes est isomorphe à l'anneau quotient* $\mathbb{R}[X]/(X^2 + 1)\mathbb{R}[X]$.

DÉMONSTRATION. — D'après (9), on a $\mathbb{C} = \mathbb{R}[i] \cong \mathbb{R}[X]/J$, où $J = \{f \in \mathbb{R}[X] \mid f(i) = 0\}$. Puisque $a + ib \neq 0$ pour $(a, b) \neq (0, 0)$ et comme $i^2 + 1 = 0 \Rightarrow X^2 + 1 \in J$, on déduit sans peine de la démonstration du corollaire au théorème 5 que $J = (X^2 + 1)\mathbb{R}[X]$.

Les éléments de l'anneau quotient $\mathbb{R}[X]/J$ sont des classes $(a + bX) + J$, avec $a, b \in \mathbb{R}$; la correspondance $a + ib \mapsto (a + bX) + J$ établit un isomorphisme de $\mathbb{C}$ sur $\mathbb{R}[X]/J$. ∎

EXERCICES

1. Les polynômes $f(X) = X^5 + 3X^4 + X^3 + 4X^2 - 3X - 1$, $g(X) = X^2 + X + 1$ peuvent être considérés comme appartenant à l'anneau $\mathbb{Z}[X]$ ou, disons, à l'anneau $Z_5[X]$, suivant l'interprétation qu'on donne à leurs coefficients. En appliquant l'algorithme de division euclidienne, montrer que dans le premier cas $f(X)$ n'est pas divisible par $g(X)$, alors que dans le deuxième cas il est divisible par $g(X)$. Une variante opposée est-elle réalisable?

2. Démontrer, en partant du théorème 3, que si F est un corps commutatif, le groupe de tous les automorphismes de l'anneau $F[X]$, identiques sur F, est isomorphe au groupe des transformations $X \mapsto aX + b$, où $a, b \in F$ et $a \neq 0$.

3. Montrer que le polynôme $f \in F[X_1, \ldots, X_n]$ est une forme de degré m (voir démonstration du théorème 4) si, et seulement si, $f(tX_1, \ldots, tX_n) = t^m f(X_1, \ldots, X_n)$, où t est une nouvelle indéterminée.

4. Montrer que le nombre de monômes distincts à n indéterminées de degré total m est égal à $\binom{m+n-1}{m}$. (Indication. Raisonner par récurrence sur n et m en se basant sur la relation suivante : $\binom{m+(n-1)-1}{m} + \binom{(m-1)+n-1}{m-1} = \binom{m+n-1}{m}$.)

5. En revenant aux définitions données au n° 1, considérons un ensemble $A[[X]]$ des *séries entières* $f(X) = \sum_{i\geqslant 0} a_i X^i$ dites *formelles* à une indéterminée X ou, si l'on veut, des suites $(a_0, a_1, a_2, \ldots)$ à n'importe quel nombre, peut-être infini, de coefficients $a_i \neq 0$ appartenant à un anneau commutatif A. Les opérations sur les séries entières formelles de $A[[X]]$ s'effectuent d'après les mêmes

règles que sur les polynômes:

$$\left(\sum a_i X^i\right)+\left(\sum b_i X^i\right)=\sum (a_i+b_i)\, X^i,$$

$$\left(\sum a_i X^i\right)\cdot\left(\sum b_j X^j\right)=\sum c_k X^k, \quad c_k=\sum_{i+j=k} a_i b_j.$$

Montrer que l'ensemble $A\,[[X]]$, muni de ces opérations, est un anneau associatif et commutatif ayant un élément unité $1 = (1, 0, 0, \ldots)$.

Puisque la série entière $f = \sum a_i X^i$ comprend des puissances X^i de l'indéterminée X avec un i aussi grand que l'on veut, il est logique de considérer, au lieu du degré $\deg f$ qui n'a maintenant aucun sens, la valuation $\omega(f)$, égale au plus petit indice n pour lequel $a_n \neq 0$ (on convient aussi que $\omega(0) = +\infty$).

Montrer que

$$\text{(i)}\ \ \omega(f-g) \geqslant \min\{\omega(f),\ \omega(g)\};\quad \text{(ii)}\ \ \omega(fg) \geqslant \omega(f)+\omega(g).$$

Si A est un anneau intègre, on a $\omega(fg) = \omega(f) + \omega(g)$. En particulier, l'anneau $A\,[[X]]$ est aussi intègre.

Montrer également que $A\,[X]$ est un sous-anneau de $A\,[[X]]$.

6. Les polynômes et les séries entières sont souvent utilisés comme *fonctions génératrices* de différentes grandeurs numériques. Expliquons sur deux exemples simples le sens des opérations sur ces fonctions.

a) Etablir la relation

$$\sum_{i=0}^{k} \binom{m}{i}\binom{n}{k-i}=\binom{m+n}{k},$$

en partant de la formule du binôme $\sum \binom{n}{i} X^i=(1+X)^n$ dans $\mathbb{Z}\,[X]$ et du développement évident $(1+X)^m(1+X)^n=(1+X)^{m+n}$.

b) Déterminer le nombre l_n de différentes possibilités de mettre les parenthèses dans un produit de n éléments d'un ensemble muni d'une opération binaire. A cet effet, il est commode d'introduire une fonction génératrice sous forme d'une série entière formelle

$$l(X)=\sum_{n\geqslant 1} l_n X^n=X+X^2+2X^3+\ldots,$$

dont les coefficients initiaux ont été calculés encore au chapitre 4, § 1, nº § 3. De la relation de récurrence évidente

$$l_n=\sum_{k=1}^{n-1} l_k l_{n-k}$$

il découle que $l(X)^2 = l(X) - X$. En résolvant cette équation du second degré on trouve

$$l(X)=\frac{1-\sqrt{1-4X}}{2}$$

(le signe du radical se détermine par la condition $l_n > 0$). Mais si la série entière $f(X)$ est telle que $f^r = 1 + \lambda X$, $r \in \mathbb{N}$, on a

$$f(X)=1+\sum_{k=1}^{\infty}\left[\prod_{i=0}^{k-1}\left(\frac{1}{r}-i\right)\right]\frac{(\lambda X)^k}{k!}$$

(c'est un développement dit « en série de Taylor » que nous demandons d'admettre pour l'instant sans démonstration). Dans notre cas on a $r = 2$, $\lambda = -4$. Une simple substitution conduit à l'expression définitive

$$l_n = \frac{1}{n}\binom{2n-2}{n-1}.$$

Nous laissons au lecteur le soin d'effectuer les calculs intermédiaires.

7. Un anneau $A\,[[X, Y]]$ des séries entières formelles à deux indéterminées indépendantes (mais permutables) X, Y a pour éléments les expressions $\sum\limits_{i\geqslant 0,\, j\geqslant 0} a_{ij}X^iY^j$. Vérifier que

$$B\,[[Y]] = A\,[[X, Y]] = C\,[[X]],$$

où $B = A\,[[X]]$, $C = A\,[[Y]]$ (réviser la construction d'un anneau des polynômes à plusieurs indéterminées). Montrer que l'intégrité de A implique celle de l'anneau $A\,[[X]]$.

§ 3. Factorisation dans l'anneau de polynômes

1. Propriétés élémentaires de la divisibilité.— Dès le chapitre 1, nous avons abordé, à de nombreuses occasions, des questions relatives à la divisibilité dans l'anneau $\mathbb{Z}$ des entiers, mais nous n'avons pas jusqu'ici démontré le théorème dit fondamental de l'arithmétique. Maintenant il est temps de non seulement combler cette lacune mais aussi d'étendre des assertions correspondantes à une plus large classe d'anneaux. C'est l'anneau $P\,[X]$ des polynômes sur un corps commutatif P qui nous intéresse en tout premier lieu.

Commençons par examiner un anneau intègre arbitraire K. Les éléments inversibles de K ont été appelés diviseurs d'unité. Souvent, on leur donne encore le nom d'éléments *réguliers*. Il est tout à fait évident qu'un polynôme $f \in A\,[X]$ est inversible (régulier) si, et seulement si, $\deg f = 0$ et $f = f_0$ est un élément inversible de l'anneau A, car $fg = 1 \Rightarrow \deg f + \deg g = \deg 1 = 0$.

On dit qu'un élément $b \in K$ *est divisible par* $a \in K$ (ou b *est un multiple de* a) s'il existe un élément $c \in K$ tel que $b = ac$ (on le note $a \mid b$). Si $a \mid b$ et $b \mid a$, les éléments a et b sont dits *associés*. Alors, $b = ua$, où $u \mid 1$. En raison de la remarque faite plus haut, la propriété des polynômes f, $g \in A\,[X]$ d'être associés signifie qu'ils ne diffèrent l'un de l'autre que par un facteur inversible appartenant à A.

Un élément $p \in K$ est dit *premier* (ou *non factorisable*) si p n'est pas inversible et ne peut pas être présenté sous la forme de $p = ab$, où a, b sont des éléments qui n'admettent pas d'inverses. Dans un corps P, tout élément non nul est inversible et il n'existe pas d'éléments premiers. Un élément premier de l'anneau $A\,[X]$ est le plus souvent appelé *polynôme premier* ou *irréductible*.

Les propriétés essentielles que présente la relation de divisibilité dans l'anneau intègre K sont :

1) *Si* $a \mid b$, $b \mid c$, *alors* $a \mid c$. En effet, on a $b = ab'$, $c = bc'$, où b', $c' \in K$. Donc, $c = (ab')\, c' = a\,(b'c')$.

2) *Si* $c \mid a$ *et* $c \mid b$, *alors* $c \mid (a \pm b)$. En effet, on a par hypothèse $a = ca'$, $b = cb'$ pour certains a', $b' \in K$ et il résulte de la distributivité que $a \pm b = c\,(a' \pm b')$.

3) *Si* $a \mid b$, *alors* $a \mid bc$. Il est clair que $b = ab' \Rightarrow bc = (ab')\, c = a\,(b'c)$.

En combinant 2) et 3), on obtient :

4) *Si* $a \in K$ *divise chacun des éléments* $b_1, b_2, \ldots, b_m \in K$, *il divisera aussi l'élément* $b_1c_1 + b_2c_2 + \ldots + b_mc_m$, *où* $c_1, c_2, \ldots, c_m$ *sont des éléments arbitraires.* ■

DÉFINITION. — *On dit qu'un anneau intègre K est un anneau factoriel (à décomposition unique en facteurs premiers) si tout élément* $a \neq 0$ *de K peut être représenté sous la forme*

$$a = up_1p_2 \ldots p_r, \tag{1}$$

où u est un élément inversible et $p_1, p_2, \ldots, p_r$ *sont des éléments premiers (non nécessairement distincts deux à deux), l'existence d'une autre décomposition* $a = vq_1q_2 \ldots q_s$ *de même type entraînant* $r = s$ *et*

$$q_1 = u_1p_1, \ldots, q_r = u_rp_r,$$

avec $u_1, \ldots, u_r$ *des éléments inversibles et une numération convenable des éléments* p_i *et* q_j.

En admettant dans l'égalité (1) que $r = 0$, nous convenons que les éléments inversibles de K admettent, eux aussi, une décomposition en facteurs premiers. Il est clair que, si p est un élément premier et u un élément inversible, l'élément up associé à p est aussi premier. Dans l'anneau $\mathbb{Z}$ à éléments inversibles 1 et -1, la relation d'ordre ($a < b$) permet de distinguer un nombre premier *positif* p parmi deux éléments premiers possibles $\pm p$. Dans l'anneau $P\,[X]$ il est commode d'examiner des polynômes irréductibles *unitaires* (c'est-à-dire ayant le coefficient dominant égal à 1).

On peut énoncer le théorème suivant :

THÉORÈME 1. — *Soit K un anneau intègre arbitraire à décomposition en facteurs premiers. Pour que l'anneau K soit factoriel, il faut et il suffit que tout élément premier* $p \in K$ *divisant le produit* $ab \in K$, *divise au moins l'un des facteurs* a, b.

DÉMONSTRATION. — Soit $ab = pc$. Etant donné

$$a = \prod a_i, \quad b = \prod b_j, \quad c = \prod c_k,$$

des décompositions de a, b, c en facteurs premiers et K, un anneau factoriel, il résulte de l'égalité $\Pi a_i \Pi b_j = p \Pi c_k$ que l'élément p est associé à l'un des a_i ou des b_j, c'est-à-dire que p divise a ou b.

Inversement, établissons l'unicité de la décomposition dans K, si $p \mid ab \Rightarrow p \mid a$ ou $p \mid b$. En raisonnant par récurrence, admettons que la décomposition de tous les éléments de K, avec un nombre de facteurs premiers $\leqslant n$, soit unique (bien sûr à l'ordre des facteurs et à des éléments associés près). Démontrons maintenant cette assertion pour tout élément $a \neq 0$ qui peut être décomposé en $n+1$ facteurs premiers. Soient

$$a = \prod_{i=1}^{n+1} p_i = \prod_{j=1}^{m+1} r_j \tag{2}$$

deux décompositions de l'élément a avec $m \geqslant n$. En appliquant l'hypothèse du théorème à $p = p_{n+1}$, nous voyons que p_{n+1} doit diviser l'un des éléments $r_1, \ldots, r_{m+1}$. Sans restreindre la généralité (puisqu'il ne s'agit que de la numération), posons $p_{n+1} \mid r_{m+1}$. Or, r_{m+1} est un élément premier et donc $r_{m+1} = up_{n+1}$, où u est un élément inversible. En utilisant la loi de simplification dans K (chap. 4, § 4, théorème 3), nous déduisons de (2) l'égalité $\prod_{i=1}^{n} p_i = u \prod_{j=1}^{m} r_j$. Le premier membre de cette égalité est le produit de n facteurs premiers. En vertu de l'hypothèse de récurrence, $m = n$ et les deux décompositions ne diffèrent l'une de l'autre que par l'ordre des facteurs premiers qui peuvent être munis de facteurs inversibles quelconques. ■

En général, dans un anneau intègre K quelconque, les éléments $a \neq 0$ peuvent ne pas admettre une décomposition de type (1). Ce qui est plus intéressant, c'est l'existence des anneaux intègres dans lesquels la décomposition en facteurs premiers, bien qu'elle soit possible, n'est pas unique; autrement dit, la condition du théorème 1, qui semble être triviale, n'est pas satisfaite dans tous les cas.

EXEMPLE. — Considérons le corps quadratique imaginaire $\mathbb{Q}(\sqrt{-5})$ (voir exemple du n° 5, § 1) et, dans ce corps, l'anneau intègre $K = \{a + b\sqrt{-5} \mid a, b \in \mathbb{Z}\}$. La norme $N(a + b\sqrt{-5}) = a^2 + 5b^2$ de tout élément $\alpha \in K$ non nul est un entier positif. Si α est inversible dans K, on a $N(\alpha)^{-1} = N(\alpha^{-1}) \in \mathbb{Z}$, d'où $N(\alpha) = 1$. Cela ne peut avoir lieu que si $b = 0$, $a = \pm 1$. Ainsi, dans K tout comme dans $\mathbb{Z}$, les seuls éléments inversibles sont ± 1. Si $\alpha = \varepsilon\alpha_1\alpha_2 \ldots \alpha_r \neq 0$, $\varepsilon = \pm 1$, alors $N(\alpha) = N(\alpha_1) \ldots N(\alpha_r)$. Puisque $1 < N(\alpha_i) \in \mathbb{N}$, le nombre r de facteurs ne peut pas croître indéfiniment pour un α donné. Il s'ensuit que la décomposition en facteurs premiers est possible dans K.

Ceci étant, le nombre 9 (et non seulement lui) admet deux décompositions en facteurs premiers essentiellement différentes:

$$9 = 3 \cdot 3 = (2 + \sqrt{-5})(2 - \sqrt{-5}).$$

Il est évident que les éléments 3 et $2 \pm \sqrt{-5}$ ne sont pas associés. En outre, $N(3) = N(2 \pm \sqrt{-5}) = 9$. Donc, il résulterait de la décomposition $\alpha = \alpha_1\alpha_2$ pour $\alpha = 3$ ou $2 \pm \sqrt{-5}$, avec α_1, α_2 non inversibles, que $9 = N(\alpha) = N(\alpha_1)N(\alpha_2)$, c'est-à-dire $N(\alpha_i) = 3$, $i = 1, 2$, ce qui est impossible, car

l'équation $x^2 + 5y^2 = 3$, $x, y \in \mathbb{Z}$, est insoluble. Cela démontre que les éléments 3 et $2 \pm \sqrt{-5}$ sont premiers.

L'exemple que nous venons de considérer, ne reflète que faiblement un grand nombre de questions relatives aux corps quadratiques $\mathbb{Q}(\sqrt{d})$, questions qui à présent ne sont résolues qu'en partie. Leur étude relève de la *théorie algébrique des nombres.*

Avant de constater, à l'aide du théorème 1, si tel ou tel anneau est factoriel, nous allons introduire des notions auxiliaires importantes qui présentent un intérêt indépendant.

2. P.G.C.D. et P.P.C.M. dans les anneaux.— Etant donné un anneau intègre K, nous appellerons *plus grand commun diviseur* de deux éléments $a, b \in K$ un élément $d \in K$ noté par le symbole P.G.C.D.(a, b) et ayant les propriétés suivantes:

(i) $d \mid a$, $d \mid b$;

(ii) $c \mid a$, $c \mid b \Rightarrow c \mid d$.

Il est clair que tout élément associé à d présente, lui aussi, les propriétés (i) et (ii). Réciproquement, si c et d sont deux plus grands diviseurs des éléments a et b, on a $c \mid d$, $d \mid c$, de sorte que c et d sont associés. Le symbole P.G.C.D.(a, b) se rapporte indifféremment à chacun d'eux, c'est-à-dire qu'avec cette écriture, les éléments associés ne se distinguent pas. Compte tenu de cette convention, on ajoutera aux propriétés (i), (ii) de la définition du plus grand commun diviseur les suivantes:

(iii) P.G.C.D.$(a, b) = a \Leftrightarrow a \mid b$;

(iv) P.G.C.D.$(a, 0) = a$;

(v) P.G.C.D.$(ta, tb) = t$ P.G.C.D.(a, b);

(vi) P.G.C.D.(P.G.C.D.(a, b), c) = P.G.C.D.(a, P.G.C.D.(b, c)).

La vérification de ces propriétés ne présentant aucune difficulté, nous laissons au lecteur le soin de la faire. La propriété (vi) permet aussi d'étendre la notion de P.G.C.D. à un nombre fini quelconque d'éléments.

Par analogie avec le P.G.C.D.(a, b) on introduit une notion duale de *plus petit commun multiple* $m =$ P.P.C.M.(a, b) des éléments $a, b \in K$, qui est aussi définie à des éléments associés près, par deux propriétés:

(i′) $a \mid m$, $b \mid m$;

(ii′) $a \mid c$, $b \mid c \Rightarrow m \mid c$.

En particulier, en posant $c = ab$, on obtient $m \mid ab$.

THÉORÈME 2. — *Soient a et b deux éléments d'un anneau intègre K, possédant un* P.G.C.D.(a, b) *et un* P.P.C.M.(a, b). *Alors:*

a) P.P.C.M.$(a, b) = 0 \Leftrightarrow a = 0$ *ou* $b = 0$;
b) $a, b \neq 0$, $m =$ P.P.C.M.(a, b), $ab = dm \Rightarrow d =$ $=$ P.G.C.D.(a, b).

Démonstration. — L'assertion a) découle directement de la définition du P.P.C.M.(a, b). Pour démontrer b) il faut nous assurer que l'élément d défini par l'égalité $ab = dm$, possède les propriétés (i), (ii). En effet, $(i') \Rightarrow m = a'a$, $m = b'b$. Par conséquent, $ab = dm = da'a$, d'où, après simplification par a admissible dans tout anneau intègre, on a $b = da'$, c'est-à-dire $d \mid b$. D'une manière analogue, $ab = dm = db'b \Rightarrow a = db'$, c'est-à-dire $d \mid a$. Nous avons retrouvé (i).

Soit maintenant $a = fa''$, $b = fb''$. Posons $c = fa''b''$. Alors $c = ab'' = ba''$ est un multiple commun à a et b. En vertu de la propriété (ii'), $c = c'm$ pour un certain $c' \in K$, d'où $fc'm = fc = f^2a''b'' = ab = dm$, c'est-à-dire $d = fc'$ et $f \mid d$. Nous avons obtenu la propriété (ii). ■

Les propriétés (i), (ii), (i'), (ii'), ainsi que le théorème 2, ne fournissent ni le mode de calcul, ni la démonstration d'existence de P.G.C.D.(a, b) et de P.P.C.M.(a, b). Le théorème 2, b) ne fait qu'établir une relation entre eux.

Supposons maintenant pour un instant que l'anneau K soit factoriel. Notons $\mathcal{P}$ un ensemble des éléments premiers de K, tel que chaque élément premier de K soit associé à un, et un seul, élément de $\mathcal{P}$. En considérant les décompositions de deux éléments $a, b \in K$, il est commode d'admettre qu'elles contiennent les mêmes éléments de $\mathcal{P}$ dont certains sont peut-être munis d'exposants zéro, c'est-à-dire

$$a = up_1^{k_1} \ldots p_r^{k_r}, \quad b = vp_1^{l_1} \ldots p_r^{l_r},$$
$$u \mid 1, \quad v \mid 1; \quad k_i \geqslant 0, \quad l_i \geqslant 0; \quad p_i \in \mathcal{P}; 1 \leqslant i \leqslant r. \tag{3}$$

En utilisant le théorème 1, on peut énoncer le caractère de divisibilité suivant, facile à retenir :

Caractère de divisibilité. — *Soient a, b deux éléments d'un anneau factoriel K, écrits sous la forme* (3). *Alors, les assertions suivantes sont vraies :*

1) $a \mid b$ *si, et seulement si,* $k_i \leqslant l_i$, $i = 1, 2, \ldots, r$;
2) P.G.C.D.$(a, b) = p_1^{s_1} \ldots p_r^{s_r}$, *où* $s_i = \min\{k_i, l_i\}$, $i = 1, 2, \ldots, r$;
3) P.P.C.M.$(a, b) = p_1^{t_1} \ldots p_r^{t_r}$, *où* $t_i = \max\{k_i, l_i\}$, $i = 1, 2, \ldots, r$. ■

Ainsi, pour s_i il faut prendre le plus petit des deux exposants k_i, l_i, et pour t_i, le plus grand de ces exposants. En particulier, les éléments $a, b \in K$ sont premiers entre eux, c'est-à-dire P.G.C.D.$(a, b) = 1$, si, et seulement si, les facteurs premiers entrant

dans la décomposition de l'un des éléments, n'entrent pas dans la décomposition de l'autre. L'inconvénient de ce caractère de divisibilité est qu'en pratique il s'avère assez difficile d'obtenir une décomposition de la forme (3). Même dans le cas où $K = \mathbb{Z}$ (ce qui ne prédétermine pas encore que $\mathbb{Z}$ est factoriel) on est amené à se contenter de légères modifications de la méthode d'essai direct des nombres premiers inférieurs au nombre n donné. Il est d'autant plus agréable qu'il existe pour les anneaux factoriels que nous allons traiter, dans ce qui suit, un procédé efficace de calcul de P.G.C.D.(a, b) et P.P.C.M.(a, b).

3. Les anneaux euclidiens sont factoriels. — La division euclidienne dans $\mathbb{Z}$ et $P[X]$ (voir chap. 1, § 8, n° 3 et chap. 5, § 2, n° 3) rend naturelle l'étude d'un anneau intègre K dans lequel à tout élément $a \neq 0$ on fait correspondre un entier non négatif $\delta(a)$, c'est-à-dire on définit l'application

$$\delta: K \setminus \{0\} = K^* \to \mathbb{N} \cup \{0\},$$

de manière à satisfaire aux conditions suivantes:

(E1) $\delta(ab) \geqslant \delta(a)$ *quels que soient* a, $b \neq 0$ *de* K;

(E2) *quels que soient* a, $b \in K$, $b \neq 0$, *il existe* q, $r \in K$ ($q =$ = « *quotient* », $r =$ « *reste* »), *tels que*

$$a = qb + r; \quad \delta(r) < \delta(b) \quad ou \quad r = 0. \tag{4}$$

Un anneau intègre K possédant ces propriétés s'appelle *anneau euclidien*. En posant $\delta(a) = |a|$ pour $a \in \mathbb{Z}$, et $\delta(a) = \deg a$ pour $a = a(X) \in P[X]$ nous arrivons à la conclusion que $\mathbb{Z}$ et $P[X]$ sont des anneaux euclidiens.

Il existe dans les anneaux euclidiens une méthode de recherche du P.G.C.D.(a, b) qui consiste en la *division suivant les puissances décroissantes* et porte le nom d'*algorithme d'Euclide*. Soient donnés deux éléments non nuls a, b d'un anneau euclidien K. En appliquant un nombre de fois suffisamment grand (mais fini) la prescription (E2), on obtient un système d'égalités de type (4), dans lequel le dernier reste est nul:

$$\begin{array}{ll} a = q_1 b + r_1, & \delta(r_1) < \delta(b), \\ b = q_2 r_1 + r_2, & \delta(r_2) < \delta(r_1), \\ r_1 = q_3 r_2 + r_3, & \delta(r_3) < \delta(r_2), \\ \dots\dots\dots\dots & \\ r_{k-2} = q_k r_{k-1} + r_k, & \delta(r_k) < \delta(r_{k-1}), \\ r_{k-1} = q_{k+1} r_k, & r_{k+1} = 0. \end{array} \tag{5}$$

Il en est réellement ainsi parce que la suite strictement décroissante d'entiers non négatifs $\delta(b) > \delta(r_1) > \delta(r_2) > \dots$ doit s'inter-

rompre, et cette interruption ne peut se produire que grâce à l'annulation de l'un des restes.

Il s avère que le dernier reste non nul r_k est justement le plus grand commun diviseur des éléments a et b au sens de la définition donnée au n° 2. En effet, par hypothèse, $r_k \mid r_{k-1}$. En se déplaçant de bas en haut dans le système (5) et en utilisant la propriété 4) de la relation de divisibilité énoncée au n° 1, on obtient une suite $r_k \mid r_{k-1}$, $r_k \mid r_{k-2}$, . . ., $r_k \mid r_2$, $r_k \mid r_1$ et, enfin, $r_k \mid b$, $r_k \mid a$. Cela signifie que r_k est diviseur commun à a et b. Réciproquement, soit c un autre diviseur quelconque des mêmes éléments. Alors on a $c \mid r_1$. En se déplaçant maintenant de haut en bas dans le système (5), on obtient une suite de relations de divisibilité $c \mid r_2$, $c \mid r_3$, ..., $c \mid r_k$ dont la dernière nous assure définitivement que le P.G.C.D.(a, b) existe et qu'on a l'égalité

$$r_k = \text{P.G.C.D.}(a, b). \tag{6}$$

Il importe ensuite d observer que chaque reste r_i dans le système (5) s'exprime par une combinaison linéaire à coefficients dans K de deux restes précédents r_{i-1} et r_{i-2}. Ceci étant, r_1 s'exprime par a et b : $r_1 = a - q_1 b$, alors que r_2, exprimé par b et r_1, représente par là même aussi une combinaison linéaire de a et b. Effectuant dans r_i les substitutions successives de r_{i-1} et r_{i-2} exprimés en fonction de a et b, on obtient pour $i = k$ l'expression

$$r_k = au + bv \tag{7}$$

faisant intervenir des éléments quelconques $u, v \in K$.

En comparant (6) et (7) et en tenant compte du théorème 2, b), on peut énoncer le théorème suivant :

Théorème 3. — *Dans un anneau euclidien K, tout couple d'éléments a, b possède un plus grand commun diviseur et un plus petit commun multiple. En se servant de l'algorithme d'Euclide on peut trouver deux éléments $u, v \in K$ tels que l'on ait la relation*

$$\text{P.G.C.D.}(a, b) = au + bv.$$

En particulier, les éléments $a, b \in K$ sont premiers entre eux si, et seulement si, il existe des éléments $u, v \in K$ tels que l'on ait

$$au + bv = 1. \blacksquare$$

Corollaire. — *Soient a, b, c des éléments d'un anneau euclidien K.*

(i) *Si* P.G.C.D.$(a, b) = 1$ *et* P.G.C.D.$(a, c) = 1$, *alors* P.G.C.D.$(a, bc) = 1$.

(ii) *Si $a \mid bc$ et* P.G.C.D.$(a, b) = 1$, *alors $a \mid c$.*

(iii) *Si $b \mid a$, $c \mid a$ et* P.G.C.D.$(b, c) = 1$, *alors $bc \mid a$.*

Démonstration. — (i) D'après le théorème 3, on a les égalités $au_1 + bv_1 = 1$, $au_2 + cv_2 = 1$. En multipliant respectivement leurs premiers et seconds membres, on trouve $a(au_1u_2 + bu_2v_1 + cu_1v_2) + bc(v_1v_2) = 1$, ce qui donne l'assertion avancée.

(ii) On a $au + bv = 1$, d'où $ac \cdot u + (bc)v = c$. Or, $bc = aw$, donc $c = a(cu + wv)$, c'est-à-dire $a \mid c$.

(iii) En vertu de la propriété (ii') du P.P.C.M., on a

$$b \mid a, \quad c \mid a \Rightarrow \text{P.P.C.M.}(b, c) \mid a \Rightarrow bc \mid a,$$

puisque $bc = \text{P.G.C.D.}(b, c) \cdot \text{P.P.C.M.}(b, c)$ et $\text{P.G.C.D.}(b, c) = 1$ par hypothèse. ∎

Le lecteur pourra étendre aisément l'assertion du théorème 3 au cas d'un nombre arbitraire fini d'éléments appartenant à un anneau euclidien.

Un pas direct en vue d'établir qu'un anneau euclidien est factoriel, est fait grâce au lemme suivant:

Lemme. — *Tout anneau euclidien K est un anneau à décomposition (c'est-à-dire tout élément $a \neq 0$ de K s'écrit sous la forme* (1)).

Démonstration. — Soit un élément $a \in K$, qui possède un diviseur propre b : $a = bc$, où b et c sont des éléments non inversibles (en d'autres termes, a et b ne sont pas associés). Démontrons que $\delta(b) < \delta(a)$.

En effet, d'après (E1) on a immédiatement $\delta(b) \leqslant \delta(bc) = \delta(a)$. Supposons vraie l'égalité $\delta(b) = \delta(a)$ et utilisons la condition (E2). On obtient q, r avec $b = qa + r$, où $\delta(r) < \delta(a)$ ou bien $r = 0$. Le cas où $r = 0$ doit être éliminé, car a et b ne sont pas associés. Pour la même raison $1 - qc \neq 0$. Par conséquent, en appliquant de nouveau (E2) (à condition de changer a en b), on a

$$\delta(a) = \delta(b) \leqslant \delta(b(1 - qc)) = \delta(b - qa) = \delta(r) < \delta(a),$$

ce qui est contradictoire. Donc, $\delta(b) < \delta(a)$.

Si maintenant $a = a_1a_2 \ldots a_n$, où tous les a_i sont non inversibles, alors $a_{m+1}a_{m+2} \ldots a_n$ est un diviseur propre de $a_ma_{m+1} \ldots a_n$, et d'après ce qui a été démontré

$$\delta(a) = \delta(a_1a_2 \ldots a_n) > \delta(a_2 \ldots a_n) > \ldots > \delta(a_n) \geqslant \delta(1).$$

Cette suite strictement décroissante d'entiers non négatifs a une longueur $n \leqslant \delta(a)$. Cela signifie qu'il existe une décomposition de a de longueur maximale, qui est une décomposition en facteurs premiers. ∎

Théorème 4. — *Tout anneau euclidien K est factoriel (autrement dit, K possède la propriété de décomposition unique en facteurs premiers).*

DÉMONSTRATION. — Compte tenu du lemme et du critère d'unicité de la décomposition énoncé dans le théorème 1, il ne reste qu'à montrer que si p est un élément premier de l'anneau K qui divise le produit bc des éléments quelconques $b, c \in K$, alors p divise b ou bien c.

En effet, pour $b = 0$ ou $c = 0$ il n'y a rien à démontrer. Si $bc \neq 0$ et $d =$ P.G.C.D.(b, p), alors d qui est diviseur de l'élément premier p, est soit égal à 1 (plus exactement, est diviseur de 1), soit associé à p. Dans le premier cas, b et p sont premiers entre eux, et l'assertion (ii) énoncée dans le corollaire du théorème 3 permet de conclure que $p \mid c$. Dans le second cas, $d = up$, $u \mid 1$ et donc $p \mid b$. ∎

COROLLAIRE. — *Les anneaux* $\mathbb{Z}$ *et* $P[X]$ *sont factoriels* (*P est un corps commutatif quelconque*).

L'unicité de la décomposition dans l'anneau de polynômes $P[X_1, \ldots, X_n]$, $n > 1$, qui n'est plus un anneau euclidien, sera établie au chapitre 9 dans lequel on trouvera d'autres exemples d'anneaux euclidiens.

4. Polynômes irréductibles. — En spécifiant la définition de l'élément premier donnée plus haut, soulignons encore une fois qu'un polynôme $f \in P[X]$ de degré non nul est dit premier ou irréductible dans $P[X]$ (ou irréductible sur le corps P) s'il n'est divisible par aucun polynôme $g \in P[X]$, tel que $0 < \deg g < \deg f$. En particulier, tout polynôme de degré 1 est irréductible. Il est tout à fait évident que l'irréductibilité d'un polynôme de degré >1 ou sa décomposition en facteurs irréductibles sont des notions qui dépendent essentiellement du corps de base P, comme le montre le polynôme $X^2 + 1 = (X + i)(X - i)$ que nous avons déjà rencontré lors de la construction du corps des nombres complexes. Ainsi, le polynôme $X^4 + 4$ est réductible sur le corps $\mathbb{Q}$ des nombres rationnels, bien que cela ne soit pas assez évident :

$$X^4 + 4 = (X^2 - 2X + 2)(X^2 + 2X + 2).$$

Les deux facteurs qui figurent au second membre de cette égalité, sont irréductibles non seulement sur $\mathbb{Q}$ mais aussi sur $\mathbb{R}$, en admettant cependant la décomposition dans $\mathbb{C}[X]$.

Tout comme dans $\mathbb{Z}$ il y a une infinité de nombres premiers (voir chap. 1, § 8), *le nombre de polynômes unitaires* (*dont le coefficient dominant est égal à* 1) *irréductibles sur un corps arbitraire P est infiniment grand.*

Dans le cas où le corps P est infini, cela est clair : il suffit de considérer les polynômes irréductibles de la forme $X - c$, $c \in P$.

Si le corps P est fini, on peut appliquer le raisonnement d'Euclide. A savoir, supposons qu'on ait déjà trouvé n polynômes irréductibles

$p_1, \ldots, p_n$. Le polynôme $f = p_1p_2 \ldots p_n + 1$ a au moins un diviseur premier unitaire, parce que $\deg f \geqslant n$. Désignons-le par p_{n+1}. Il est distinct de $p_1, \ldots, p_n$, sinon $p_{n+1} = p_s$ pour un $s \leqslant n$ quelconque entraînerait $p_s \mid (f - p_1 \ldots p_n)$, c'est-à-dire $p_s \mid 1$. ■

Puisque le nombre de polynômes de degré donné sur un corps fini est fini, on peut faire la conclusion utile suivante:

Sur tout corps fini il existe des polynômes irréductibles de degré aussi élevé que l'on veut. ■

Cette assertion de nature qualitative sera précisée dans le chapitre 9.

Les polynômes irréductibles sur le corps $\mathbb{Q}$ sont d'une importance particulière dans la théorie des corps des nombres algébriques. La multiplication par un entier naturel convenable permettant toujours d'effectuer le passage d'un polynôme de $\mathbb{Q}[X]$ à un polynôme de $\mathbb{Z}[X]$, il est logique de préciser d'abord la relation qui existe entre les propriétés de la réductibilité sur $\mathbb{Q}$ et sur $\mathbb{Z}$. Ayant en vue d'autres applications, nous allons démontrer une proposition générale relative aux polynômes sur un anneau factoriel K. Associons à tout polynôme $f = a_0 + a_1X + \ldots + a_nX^n \in K[X]$ le plus grand commun diviseur $\mathrm{d} = \mathrm{d}(f)$ de tous ses coefficients. Jusqu'ici, nous avons opéré avec le P.G.C.D.(a, b) de deux éléments, mais les propriétés (i) à (vi) du P.G.C.D. permettent d'étendre aisément cette notion à n'importe quel nombre fini d'éléments d'un anneau intègre. Si $\mathrm{d}(f)$ est un élément inversible de K, on dit que le polynôme f est *primitif*.

LEMME DE GAUSS. — *Soient K un anneau factoriel et $f, g \in K[X]$. Alors*

$$\mathrm{d}(fg) \approx \mathrm{d}(f)\cdot\mathrm{d}(g).$$

En particulier, si deux polynômes sont primitifs il en sera de même de leur produit (ici et plus loin on considère les égalités à l'association près, ce qui veut dire que $\mathrm{d}(fg)$ *et* $\mathrm{d}(f)\cdot\mathrm{d}(g)$ *sont des éléments associés).*

DÉMONSTRATION. — Commençons par la dernière assertion. Soient

$$f = a_0 + a_1X + \ldots + a_nX^n, \quad g = b_0 + b_1X + \ldots + b_mX^m$$

deux polynômes primitifs à coefficients dans K dont le produit fg n'est pas primitif. Cela signifie qu'il existe un élément premier $p \in K$ qui divise $\mathrm{d}(fg)$. Choisissons les plus petits indices s, t tels que $p \nmid a_s$, $p \nmid b_t$. De tels indices existent du fait que f et g sont primitifs. Le coefficient de X^{s+t} dans fg sera

$$c_{s+t} = a_sb_t + (a_{s+1}b_{t-1} + a_{s+2}b_{t-2} + \ldots) + \\ + (a_{s-1}b_{t+1} + a_{s-2}b_{t+2} + \ldots).$$

Puisque a_{s-i} et b_{t-i}, $i > 0$, sont divisibles par p conformément à l'énoncé, et $p \mid c_{s+t}$ en vertu de l'hypothèse faite, on a la relation

$$pu = a_s b_t + pv,$$

dont il ressort que $p \mid a_s b_t$. L'anneau K étant factoriel, on a $p \mid a_s$ ou $p \mid b_t$, ce qui est contradictoire. L'assertion est donc démontrée.

En passant au cas général, écrivons des polynômes quelconques f, $g \in K[X]$ sous la forme

$$f = \mathrm{d}(f) f_0, \quad g = \mathrm{d}(g) g_0,$$

où f_0, g_0 sont des polynômes primitifs. Puisque $fg = \mathrm{d}(f)\,\mathrm{d}(g)\cdot f_0 g_0$ et, d'après ce qui vient d'être démontré, $\mathrm{d}(f_0 g_0) \approx 1$, on a $\mathrm{d}(fg) \approx$ $\approx \mathrm{d}(f)\,\mathrm{d}(g)$. ■

Corollaire. — *Si un polynôme $f \in \mathbb{Z}[X]$ est irréductible sur $\mathbb{Z}$, il l'est aussi sur $\mathbb{Q}$* ($\deg f > 0$).

Démonstration. — D'après le corollaire du théorème 4, $\mathbb{Z}$ est un anneau factoriel, et donc le lemme de Gauss est applicable à $\mathbb{Z}[X]$. Supposons $f = gh$, où $f \in \mathbb{Z}[X]$ et $g, h \in \mathbb{Q}[X]$. En multipliant les deux membres de cette égalité par le plus petit commun multiple des dénominateurs de tous les coefficients de g et h, nous la mettrons sous la forme $af = bg_0h_0$, où $a, b \in \mathbb{Z}$ et g_0, h_0 sont des polynômes primitifs sur $\mathbb{Z}$. Suivant le lemme de Gauss on a $a\,\mathrm{d}(f) =$ $= b$ (dans ce cas, il est possible, sans restreindre la généralité, de remplacer $\approx$ par $=$), si bien qu'on obtient une décomposition $f =$ $= \mathrm{d}(f)\, g_0 h_0$ dans $\mathbb{Z}[X]$. Il ne reste qu'à se rappeler que f est irréductible dans $\mathbb{Z}[X]$. ■

Critère d'irréductibilité (Eisenstein). — *Soit*

$$f(X) = X^n + a_1X^{n-1} + \ldots + a_{n-1}X + a_n$$

un polynôme unitaire sur $\mathbb{Z}$ dont tous les coefficients $a_1, \ldots, a_n$ sont divisibles par un certain nombre premier p, et a_n n'est pas divisible par p^2. Alors, $f(X)$ est irréductible sur $\mathbb{Q}$.

En effet, en raisonnant par l'absurde et en utilisant le corollaire du lemme de Gauss, écrivons f sous la forme du produit de deux polynômes unitaires sur $\mathbb{Z}$:

$$f(X) = (X^s + b_1X^{s-1} + \ldots + b_s)(X^t + c_1X^{t-1} + \ldots + c_t),$$
$$st > 0.$$

Cette décomposition restera valable aussi dans l'anneau quotient $\mathbb{Z}[X]/(p) \cong Z_p[X]$ dont les éléments sont obtenus à partir des polynômes sur $\mathbb{Z}$, en prenant leurs coefficients par rapport au module p. Par hypothèse, on a $\bar{a}_i = \bar{0}$, où $\bar{a}_i$ est une classe résiduelle modulo p, correspondant à l'entier a_i. Or, l'anneau $Z_p[X]$ est fac-

toriel (corollaire du théorème 4). En comparant les deux décompositions

$$X^s X^t = (X^s + \bar{b}_1 X^{s-1} + \ldots)(X^t + \bar{c}_1 X^{t-1} + \ldots), \quad s + t = n,$$

nous arrivons inévitablement à cette conclusion que $\bar{b}_i = \bar{0} = \bar{c}_j$, c'est-à-dire que tous les coefficients b_i, c_j sont divisibles par p. Dans ce cas, $a_n = b_s c_t$ est divisible par p^2 : contradiction qui démontre le critère d'Eisenstein. ■

Remarque. — Le critère d'Eisenstein reste valable dans le cas où le coefficient dominant a_0 est différent de 1 et n'est pas divisible par p.

Exemple.— *Le polynôme* $f(X) = X^{p-1} + X^{p-2} + \ldots + X + 1$ *est irréductible sur* $\mathbb{Q}$ *quel que soit le nombre premier* p.

Il suffit d'observer que la propriété de $f(X)$ d'être irréductible est équivalente à celle du polynôme

$$f(X+1) = \frac{(X+1)^p - 1}{(X+1)-1} = X^{p-1} + \binom{p}{1} X^{p-2} + \ldots + \binom{p}{p-2} X + \binom{p}{p-1},$$

dont tous les coefficients, sauf le premier, sont divisibles par p (propriété des coefficients binomiaux que nous avons indiquée dans le chap. 4, § 4, exercice 8), et auquel on peut donc appliquer le critère d'Eisenstein. ■

EXERCICES

1. Montrer que

$$n\mathbb{Z} + m\mathbb{Z} = \mathbb{Z}\cdot\text{P.G.C.D.}(n, m),$$
$$n\mathbb{Z} \cap m\mathbb{Z} = \mathbb{Z}\cdot\text{P.P.C.M.}(n, m).$$

2. Soient f, g deux polynômes unitaires de $\mathbb{Z}[X]$. Montrer que dans l'expression P.G.C.D. $(f, g) = fu + gv$, avec $u, v \in \mathbb{Z}[X]$, on peut considérer $\deg u < \deg g$, $\deg v < \deg f$.

3. Les anneaux $\mathbb{Z}[\sqrt{-3}]$ et $Z_8[X]$ sont-ils factoriels?

4. Décomposer en facteurs irréductibles dans $\mathbb{Z}[X]$ les polynômes $X^n - 1$, $5 \leqslant n \leqslant 12$.

5. Démontrer que les facteurs irréductibles du polynôme homogène

$$f(X, Y) = a_0 X^n + a_1 X^{n-1} Y + \ldots + a_{n-1} X Y^{n-1} + a_n Y^n \in \mathbb{Q}[X, Y]$$

sont homogènes, et $f(X, Y)$ est irréductible si, et seulement si, le polynôme $f(X, 1) = a_0 X^n + a_1 X^{n-1} + \ldots + a_{n-1} X + a_n \in \mathbb{Q}[X]$ est irréductible.

6. Soient P un corps commutatif et $f(X) = \sum_{i \geqslant 0} a_i X^i$ une série entière formelle de $P[[X]]$ (voir exercice 5 du § 2). La condition $a_0 \neq 0$ ou, ce qui est équivalent, $\omega(f) = 0$ est une condition nécessaire et suffisante pour qu'existe une série entière $g(X) \in P[[X]]$ inverse de f : $fg = 1$. Par exemple, $(1 - X)^{-1} = \sum_{i \geqslant 0} X^i$. X est le seul, à l'association près, élément premier de $P[[X]]$. L'anneau $P[[X]]$ est factoriel. Justifier toutes ces assertions.

7. Montrer que $\det(x_{ij}) = \sum_{\pi \in S_n} \varepsilon_\pi x_{\pi(1), 1} \cdots x_{\pi(n), n}$ est un polynôme homogène irréductible de degré n à n^2 indéterminées indépendantes x_{ij}. (I n d i c a t i o n. En raisonnant par l'absurde, supposer que $\det(x_{ij}) =$

$= g_1(\ldots, x_{ij}, \ldots) g_2(\ldots, x_{ij}, \ldots)$. Puisque $\det(x_{ij})$ est un polynôme homogène linéaire à indéterminées se trouvant dans une seule colonne fixe, l'un des facteurs g_1, g_2 est un polynôme homogène linéaire en x_{ij}, $1 \leqslant i \leqslant n$, pour un j fixe, alors que l'autre est tout à fait indépendant de x_{ij}, $1 \leqslant i \leqslant n$. Les raisonnements analogues restent valables si l'on remplace les colonnes par les lignes. Supposons, par exemple, que x_{11} appartienne à g_1. Alors, g_2 ne contient pas x_{1j}, $1 \leqslant j \leqslant n$, d'où il résulte que g_2 ne contient pas x_{ij}, $1 \leqslant i, j \leqslant n$, ce qui signifie que g_2 est une constante.)

§ 4. Corps des quotients

1. Construction du corps des quotients d'un anneau intègre.— Au cours de deux paragraphes précédents nous avons établi beaucoup de propriétés communes à $\mathbb{Z}$ et à $P[X]$. Maintenant, notre but immédiat est d'inclure $P[X]$ dans un corps et de le faire par le procédé le plus économique, comme par exemple l'inclusion de $\mathbb{Z}$ dans $\mathbb{Q}$. Au fait, le problème posé ne serait pas plus délicat s'il s'agissait d'un anneau intègre arbitraire A.

Considérons l'ensemble $A \times A^*$ $(A^* = A \setminus \{0\})$ de tous les couples (a, b) des éléments $a, b \in A$, avec $b \neq 0$. Faisons une partition de cet ensemble en classes, en posant que les couples (a, b) et (c, d) appartiennent à une seule et même classe dès que $ad = bc$, ce qui est noté $(a, b) \sim (c, d)$. Il est clair que l'on a toujours $(a, b) \sim (a, b)$. On a aussi $(a, b) \sim (c, d) \Leftrightarrow (c, d) \sim (a, b)$ et $(a, b) \sim (c, d)$, $(c, d) \sim (e, f) \Rightarrow (a, b) \sim (e, f)$. En effet, on a les égalités $ad = bc$, $cf = de$, d'où $adf = bcf = bde$, ou bien $d(af - be) = 0$. Or, $d \neq 0$ par hypothèse. L'anneau A étant intègre, on obtient $af = be$, ce qui signifie justement que $(a, b) \sim (e, f)$. Ainsi, la relation $\sim$ est réflexive, symétrique et transitive, c'est-à-dire elle est (voir chap. 1, § 6) une relation d'équivalence dans l'ensemble $A \times A^*$ et définit donc la partition de cet ensemble en classes disjointes.

Soit $Q(A)$ l'ensemble de toutes les classes d'équivalence ou, ce qui revient au même, $Q(A)$ est un ensemble quotient $A \times A^*/\sim$ de l'ensemble $A \times A^*$ par la relation d'équivalence $\sim$. Nous désignons par le symbole $[a, b]$ la classe qui comprend le couple ordonné (a, b). Par définition,

$$[a, b] = [c, d] \Leftrightarrow ad = bc. \qquad (1)$$

Si, dans l'ensemble $A \times A^*$, on définit les opérations d'addition et de multiplication par les formules

$$(a, b) + (c, d) = (ad + bc, bd); \quad (a, b)(c, d) = (ac, bd)$$

(cela est possible parce que, dans A, $b \neq 0$, $d \neq 0$ impliquent $bd \neq 0$), ces opérations binaires peuvent être appliquées à $Q(A)$.

En effet, il nous faut montrer que

$$(a', b') \sim (a, b) \Rightarrow \begin{cases} (a, b)+(c, d) \sim (a', b')+(c, d), \\ (a, b)\cdot(c, d) \sim (a', b')\cdot(c, d). \end{cases}$$

La même chose s'exprime par les relations

$$(ad + bc)\, b'd = (a'd + b'c)\, bd,$$
$$ac \cdot b'd = a'c \cdot bd,$$

dont la vérité découle immédiatement de la condition $a'b = ab'$. Le résultat est analogue si l'on remplace (c, d) par (c', d'), où $cd' = c'd$. On peut conclure que

$$[a, b] + [c, d] = [ad + bc, bd]\,; \quad [a, b]\,[c, d] = [ac, bd] \qquad (2)$$

définissent dans $Q(A)$ les opérations d'addition et de multiplication, indépendamment des représentants choisis dans les classes d'équivalence. Ici, il faudrait mieux écrire $[a, b] \oplus [c, d]$ et $[a, b] \odot [c, d]$, mais, sans nuire à la clarté, les signes $\oplus$ et $\odot$ sont remplacés par les signes habituels d'addition et de multiplication.

Vérifions maintenant que l'ensemble $Q(A)$ muni des opérations (2) est un corps commutatif. En effet, l'associativité de l'addition découle, par exemple, des relations

$$\begin{aligned} [a, b] + ([c, d] + [e, f]) &= [a, b] + [cf + de, df] = \\ &= [adf + bcf + bde, bdf], \\ ([a, b] + [c, d]) + [e, f] &= [ad + bc, bd] + [e, f] = \\ &= [adf + bcf + bde, bdf]. \end{aligned}$$

L'associativité de la multiplication est évidente. Les relations

$$\begin{aligned} ([a, b] + [c, d])\cdot[e, f] &= [ade + bce, bdf], \\ [a, b]\,[e, f] + [c, d]\,[e, f] &= [adef + bcef, bfdf] = \\ &= [(ade + bce)\, f, (bdf)\, f] \end{aligned}$$

et la condition (1) d'égalité des classes d'équivalence montrent que la loi de distributivité est, elle aussi, remplie. On vérifie aussi sans peine la commutativité des opérations d'addition et de multiplication. L'élément neutre pour l'addition est $[0, 1]$ ($[0, 1] + [a, b] = [a, b]$), et $[1, 1]$ pour la multiplication. On a aussi $-[a, b] = [-a, b]$ puisque $[a, b] + [-a, b] = [0, b^2] = [0, 1]$. Tout cela, pris ensemble, signifie que $Q(A)$ est un anneau unitaire commutatif. Si $[a, b] \neq [0, 1]$, alors $a \neq 0$ dans A, donc $[b, a] \in Q(A)$ et $[a, b]\,[b, a] = [1, 1]$, de sorte que l'inverse de $[a, b] \neq [0, 1]$ est $[b, a]$. Nous avons démontré par là même que $Q(A)$ est un corps commutatif.

La correspondance $a \mapsto [a, 1]$ définit une application injective $f: A \to Q(A)$ qui est en fait un (mono)morphisme des anneaux ($f(a+b) = f(a) + f(b)$, $f(ab) = f(a) f(b)$; $a \neq b \Rightarrow f(a) \neq \neq f(b)$). Pour tout élément $x = [a, b] \in Q(A)$ on a

$$[b, 1]\, x = [a, 1],$$

de sorte que x est le « quotient » $f(a)/f(b)$ des éléments de $f(A)$. C'est la raison pour laquelle $Q(A)$ s'appelle *corps des quotients* de l'anneau A.

Il est commode d'identifier chaque élément $a \in A$ à son image $f(a) = [a, 1] \in Q(A)$, c'est-à-dire de remplacer A par $f(A)$. On peut faire autrement: remplacer chacun des éléments $[a, 1] \in Q(A)$ par $a \in A$, en conservant tous les autres éléments du corps $Q(A)$, et effectuer des substitutions correspondantes dans les formules (2). A savoir, il convient de poser

$$a + [b, c] = [ac + b, c]; \quad a[b, c] = [ab, c].$$

Il en résulte que l'anneau intègre A sera dès le début un sous-anneau d'un corps isomorphe à $Q(A)$ et représenté généralement par le même symbole $Q(A)$. Après une telle convention il est logique de donner aux éléments $[a, b]$ le nom de *fractions* et de les écrire sous la forme habituelle plus courte

$$[a, b] = \frac{a}{b}.$$

Les opérations sur les classes $[a, b]$, introduites plus haut, répètent, comme il est facile de s'en apercevoir, les opérations sur les fractions dans un corps commutatif (voir chap. 4, § 4, n° 5, (10)). Nous avons ainsi démontré le théorème suivant:

THÉORÈME 1. — *Pour tout anneau intègre A il existe un corps des quotients* (*ou corps des fractions*) $Q(A)$ *dont les éléments sont de la forme* a/b, $a \in A$, $0 \neq b \in A$. *Les opérations sur les fractions obéissent aux règles* (1), (2) *dans lesquelles il convient de poser* $[a, b] = a/b$. ■

La construction de corps des quotients est assez fréquemment utilisée en mathématiques. Son naturel se justifie ne serait-ce par le fait que le corps $\mathbb{Q}$ n'est rien d'autre que le corps des quotients $\mathbb{Q}(\mathbb{Z})$ de l'anneau $\mathbb{Z}$. Il est facile de voir (vérifiez-le!) que $Q(A) \cong \cong A$ si A est un corps.

REMARQUE. — On peut démontrer que, si un anneau intègre A est un sous-anneau d'un corps P dans lequel chaque élément x s'écrit sous la forme du quotient a/b des éléments $a \in A$, $0 \neq b \in A$, alors $P \cong Q(A)$ Par exemple, $\mathbb{Q}(\sqrt{d}) \cong Q(\mathbb{Z}[\sqrt{d}])$.

2. Corps des fractions rationnelles.— Soit P un corps commutatif et soit $P[X]$ un anneau des polynômes à coefficients dans P. Le

corps des quotients $Q(P[X])$ de l'anneau $P[X]$ se note $P(X)$ (on remplace les crochets par les parenthèses) et s'appelle *corps des fractions rationnelles* à une indéterminée X et à coefficients dans P.

Il importe de remarquer que le corps des fractions rationnelles $P(X)$ contient toujours un nombre infini d'éléments et que sa caractéristique coïncide avec celle du corps P. Le corps $\mathbb{F}_p(X)$ fournit un exemple de corps infini de caractéristique $p > 0$.

Toute fraction rationnelle du corps $P(X)$ s'écrit (et de plusieurs manières) sous la forme f/g (ou $\frac{f}{g}$ si l'on ne cherche pas à économiser sur le papier), où f, g sont des polynômes de l'anneau $P[X]$, $g \neq 0$. Par définition, $f/g = f_1/g_1 \Leftrightarrow fg_1 = f_1g$. Il est naturel d'appeler f *numérateur* et g *dénominateur de la fraction* f/g. La fraction ne change pas si son numérateur et son dénominateur sont multipliés par un même polynôme non nul ou simplifiés par un facteur commun quelconque. En particulier, l'entier (positif ou négatif) $\deg f - \deg g$ ne dépend pas de la représentation de la fraction rationnelle non nulle sous forme de rapport (de quotient) f/g de deux polynômes. Ce nombre s'appelle *degré de la fraction*. Une fraction rationnelle à une indéterminée X est dite *irréductible* si son numérateur et son dénominateur sont premiers entre eux. Toute fraction rationnelle f/g est définie d'une façon unique, à un facteur de P, commun au numérateur et au dénominateur, près, par une fraction irréductible. En effet, la division de f et g par le P.G.C.D.(f, g) aboutit à une fraction irréductible. Or, l'égalité $f/g = f_1/g_1$ de deux fractions irréductibles exprimée sous la forme $fg_1 = f_1g$ donne $f = cf_1$, $c \in P$, $g = cg_1$ (utiliser le corollaire au théorème 4 du § 3).

Si $\deg(f/g) = \deg f - \deg g < 0$, on dit que la fraction (irréductible) f/g est *propre* (le polynôme nul est une fraction propre parce que nous avons convenu de considérer $\deg 0 = -\infty$).

THÉORÈME 2. — *Toute fraction rationnelle de $P(X)$ peut se représenter d'une manière et d'une seule sous forme de la somme d'un polynôme et d'une fraction propre.*

DÉMONSTRATION. — L'algorithme de division euclidienne appliqué au numérateur et au dénominateur de la fraction f/g donne l'égalité $f = qg + r$, où $\deg r < \deg g$. Maintenant, $f/g = q + r/g$ est l'écriture cherchée dont la comparaison avec toute autre écriture de même type $f/g = \bar{q} + \bar{r}/\bar{g}$ ($\bar{q}$, $\bar{r}$, $\bar{g} \in P[X]$, $\deg \bar{r} < \deg \bar{g}$) conduit à la relation

$$\bar{q} - q = r/g - \bar{r}/\bar{g} = (r\bar{g} - \bar{r}g)/g\bar{g}.$$

Puisque $\bar{q} - q \in P[X]$, et

$$\deg((r\bar{g} - \bar{r}g)/g\bar{g}) = \deg(r\bar{g} - \bar{r}g) - \deg g - \deg \bar{g} < 0,$$

cela ne peut avoir lieu que dans le cas où $q - \bar{q} = 0$ et $r/g = \bar{r}/\bar{g}$. ■

3. Fractions simples.— On dit qu'une fraction rationnelle propre $f/g \in P(X)$ est *simple* si $g = p^n$, $n \geqslant 1$, où $p = p(X)$ est un polynôme irréductible et $\deg f < \deg p$.

Le théorème fondamental relatif aux fractions rationnelles est le suivant:

THÉORÈME 3. — *Toute fraction rationnelle propre peut se décomposer d'une manière et d'une seule en somme de fractions simples.*

DÉMONSTRATION.— Elle se subdivise en deux parties: la démonstration de l'existence de la décomposition et celle de son unicité.

I. Soit donnée une fraction rationnelle propre $f/g \in P(X)$ dans laquelle le polynôme g peut, sans restreindre la généralité, être considéré comme unitaire. Supposons que $g = g_1g_2$ soit le produit de deux polynômes unitaires premiers entre eux. Conformément aux résultats du § 3, on a la relation

$$1 = u_1g_1 + u_2g_2$$

avec certains u_1, $u_2 \in P[X]$. La multiplication par f des deux membres de cette égalité donne

$$f = fu_1g_1 + fu_2g_2.$$

Si $fu_1 = qg_2 + v_2$, $\deg v_2 < \deg g_2$, alors

$$f = v_1g_2 + v_2g_1, \tag{3}$$

où $v_1 = qg_1 + fu_2$. La fraction étant propre, on a $\deg f < \deg g$, si bien que la relation (3) ne peut avoir lieu que pour $\deg v_1 < \deg g_1$, car $\deg v_2 < \deg g_2$.

En divisant les deux membres de la relation (3) par g_1g_2, on obtient la décomposition de la fraction f/g en somme de deux fractions:

$$f/g = v_1/g_1 + v_2/g_2,$$

ceci étant, les deux fractions figurant au second membre de cette égalité sont propres. Si le dénominateur g_i de l'une de ces fractions se représente de nouveau par le produit de deux polynômes premiers entre eux, on peut appliquer à cette fraction les raisonnements précédents et obtenir donc pour elle une décomposition de même type. En opérant de cette façon, on obtient finalement une somme

$$\frac{f}{g} = \sum_{i=1}^{m} \frac{a_i}{p_i^{n_i}}, \tag{4}$$

P.G.C.D.$(a_i, p_i) = 1$, $\deg a_i < n_i \deg p_i$, où les dénominateurs sont les puissances $p_i^{n_i}$ des polynômes unitaires irréductibles p_i

dans la décomposition de g:

$$g = p_1^{n_1} p_2^{n_2} \dots p_m^{n_m} \tag{5}$$

($p_i \neq p_j$ pour $i \neq j$).

Décomposons plus loin la fraction propre a/p^n. Puisque par hypothèse $\deg a < n \deg p$, l'algorithme de division euclidienne nous amène à un système d'égalités

$$\begin{aligned} a &= q_1 p^{n-1} + r_1, \\ r_1 &= q_2 p^{n-2} + r_2, \\ &\dots\dots\dots \\ r_{n-2} &= q_{n-1} p + r_{n-1}, \\ r_{n-1} &= q_n, \end{aligned}$$

où $\deg q_i < \deg p$ pour tous les quotients $q_1, \dots, q_n$. On voit que

$$a = q_1 p^{n-1} + q_2 p^{n-2} + \dots + q_{n-1} p + q_n,$$

d'où

$$\frac{a}{p^n} = \frac{q_1}{p} + \frac{q_2}{p^2} + \dots + \frac{q_{n-1}}{p^{n-1}} + \frac{q_n}{p^n}.$$

Puisque $\deg q_i < \deg p$, les fractions q_i/p^i sont simples.

II. En passant à l'unicité de la décomposition, supposons qu'en plus de la représentation

$$\frac{f}{g} = \sum_{i=1}^{m} \left(\sum_{j=1}^{n_i} \frac{a_{ij}}{p_i^j} \right), \quad \deg a_{ij} < \deg p_i \tag{6}$$

obtenue pour la fraction propre f/g sous la forme de somme des fractions simples, il existe encore une décomposition

$$\frac{f}{g} = \sum_{k=1}^{\mu} \left(\sum_{l=1}^{\nu_i} \frac{b_{kl}}{q_k^l} \right),$$

dans laquelle peuvent se rencontrer des termes b_{kl}/q_k^l dont les dénominateurs q_k^l ne sont pas dans (6). Introduisons, s'il y a lieu, dans les deux expressions de f/g les termes à numérateurs nuls a_{ij} et b_{kl} et retranchons membre à membre une expression de l'autre. En réduisant les termes semblables (ayant les mêmes dénominateurs), on obtient l'identité

$$\sum_{i=1}^{M} \left(\sum_{j=1}^{N_i} \frac{a_{ij} - b_{ij}}{p_i^j} \right) = 0. \tag{7}$$

Ici, $M \leqslant m + \mu$, p_i, $i > m$, désigne un élément q_k quelconque, alors que les N_i sont choisis de manière que

$$a_{i, N_i} - b_{i, N_i} \neq 0. \tag{8}$$

En multipliant (7) par $\prod_{i=1}^{M} p_i^{N_i}$, on obtient l'identité polynomiale

$$(a_{1,N_1}-b_{1,N_1})\prod_{i=2}^{M} p_i^{N_i}+p_1u=0.$$

La forme du polynôme u ne nous intéresse pas. Il importe que cette identité entraîne la divisibilité de $(a_{1,N_1}-b_{1,N_1})\prod_{i=2}^{M} p_i^{N_i}$ par p_1. Or, P.G.C.D.$(\prod_{i=2}^{M} p_i^{N_i},\ p_1)=1$, donc $p_1 \mid (a_{1,N_1}-b_{1,N_1})$. Il reste à se rappeler que $\deg (a_{1,N_1}-b_{1,N_1}) \leqslant \max \{\deg a_{1,N_1}, \deg b_{1,N_1}\} < \deg p_1$. Par conséquent $a_{1,N_1}-b_{1,N_1}=0$: contradiction avec l'hypothèse (8). ■

La démonstration du théorème 3 est parfaitement constructive (si l'on suppose connue la décomposition (4)) et peut être utilisée en pratique pour représenter une fraction propre sous la forme d'une somme des fractions simples.

Remarquons que si $g=(X-c)^n h$, $h(c)\neq 0$, on a

$$\frac{f}{g}=\frac{b_1}{(X-c)^n}+\frac{f-b_1h}{(X-c)^n h}$$

pour tout $b_1 \in P$. En posant $b_1=f(c)/h(c)$, on obtient $f(c)-b_1h(c)=0$ et par conséquent $f-b_1h=(X-c)f_1$ (voir chap. 5, § 1 et chap. 6). Ainsi

$$\frac{f-b_1h}{(X-c)^n h}=\frac{f_1}{(X-c)^{n-1} h}.$$

En appliquant à cette fraction le même procédé, nous réduisons encore d'une unité l'exposant de $X-c$ au dénominateur et ainsi de suite. Après n pas, nous obtiendrons la décomposition

$$\frac{f}{g}=\frac{f}{(X-c)^n h}=\frac{f_0}{h}+\sum_{i=1}^{n}\frac{b_{1+n-i}}{(X-c)^i},\quad b_k\in P.$$

Si h (donc $g=(X-c)^n h$) se décompose entièrement en facteurs linéaires, nous pouvons, en détachant l'une après l'autre les fractions simples $b_{ki}/(X-c_i)^k$ et en utilisant la propriété d'unicité de la décomposition, venir par une voie légèrement différente au résultat énoncé dans le théorème.

Dans le cas où tous les facteurs irréductibles p_i dans (4) sont linéaires ou quadratiques et par conséquent les fractions simples sont de la forme

$$\frac{d}{(X-c)^n}\quad\text{ou}\quad\frac{dX+e}{(X^2+aX+b)^n};\quad a, b, c, d, e\in P, \tag{9}$$

il est aussi commode d'appliquer une méthode appelée *méthode des coefficients indéterminés.* Elle se résume ainsi : on écrit f/g sous la forme d'une somme des fractions de type (9), on multiplie les deux membres de l'égalité par g et dans la relation obtenue pour les polynômes, on donne à X des valeurs convenables de P, afin de déterminer les coefficients d, e, . . . Les résultats qui seront obtenus au chapitre suivant, montreront que la méthode des coefficients indéterminés est applicable sans restriction si P est le corps des nombres complexes ou des nombres réels. Or, c'est justement sur ces corps qu'on considère le plus souvent les fractions simples dont on se sert comme d'un instrument technique lors de l'intégration des fonctions rationnelles.

EXERCICES

1. Construire le corps des quotients $\mathbb{R}((X))$ de l'anneau $\mathbb{R}[[X]]$ des séries entières formelles à une indéterminée X à coefficients dans $\mathbb{R}$. En se servant des résultats obtenus dans l'exercice 6 du § 3, montrer que chaque élément du corps $\mathbb{R}((X))$ a la forme d'une *série entière* dite *méromorphe*

$$\varphi(X) = a_{-m}X^{-m} + a_{-m+1}X^{-m+1} + \ldots$$
$$\ldots + a_{-1}X^{-1} + a_0 + a_1X + a_2X^2 + \ldots, \quad a_i \in \mathbb{R},$$

qui admet un nombre fini d'exposants négatifs. En d'autres termes, $\varphi(X) = X^{-m}f(X)$, où $f(X)$ est une série entière ordinaire de $\mathbb{R}[[X]]$.

2. Entendons par $\mathbb{R}(X, Y)$ (respectivement par $\mathbb{R}((X, Y))$) le corps des quotients de l'anneau des polynômes $\mathbb{R}[X, Y]$ (respectivement de l'anneau intègre $\mathbb{R}[[X, Y]]$; voir exercice 7 du § 2). Montrer que

$$\mathbb{R}(X, Y) = \{\mathbb{R}(X)\}(Y) = \{\mathbb{R}[X]\}(Y).$$

Les corps $\mathbb{R}((X, Y))$ et $\{\mathbb{R}((X))\}((Y))$ sont-ils isomorphes? (R é p o n s e. Non).

3. Soit a_0, a_1, a_2, . . . une suite infinie de nombres réels, périodique à partir d'un certain terme. Montrer que la série entière $f(X) = a_0 + a_1X + a_2X^2 + \ldots$ s'écrit sous forme d'une fraction rationnelle de $\mathbb{R}(X)$.

4. Soient K un anneau commutatif ayant un élément unité 1, non nécessairement intègre, M un sous-monoïde du monoïde multiplicatif dans K, $S = K \times M$. Montrer que la relation binaire de graphe $\Gamma \subset S^2$ définie par la condition

$$\Gamma = \{((a, b), (c, d)) \in S^2 \mid (ad - bc)u = 0, \quad u \in M\}$$

(on a en vue un certain $u \in M$) est une relation d'équivalence dans S. (I n d i c a t i o n. La réflexivité et la symétrie de la relation sont évidentes. Si maintenant $((a, b), (c, d)) \in \Gamma$ et $((c, d), (e, f)) \in \Gamma$, si bien que $(ad - bc)u = 0$ et $(cf - de)v = 0$ pour certains $u, v \in M$, alors il convient de multiplier la première égalité par fv et la deuxième par bu. Leur addition membre à membre donnera l'égalité $(af - be)duv = 0$, avec $duv \in M$, puisque $d, u, v \in M$ et M est un monoïde. Donc, $((a, b), (e, f)) \in \Gamma$ et la transitivité est démontrée.

5. En copiant la démonstration du théorème 1, montrer que sur l'ensemble quotient S/Γ de l'ensemble $S = K \times M$ par la relation d'équivalence de l'exercice 4, on peut introduire une structure d'anneau unitaire commutatif. Cet anneau $Q_M(K)$ s'appelle *anneau des quotients sur K par rapport à M*. Pour un anneau intègre K et pour $M = K^*$ on obtient le corps des quotients ordinaire

$Q(K)$. (Indication. Soit $a/b = [a, b]$ une classe d'équivalence de représentant $(a, b) \in S$. Introduire deux opérations binaires $\oplus$ $\odot$:

$$a/b \oplus c/d = (ad + bc)/bd, \qquad a/b \odot c/d = ac/bd$$

et montrer que cette définition ne dépend pas du choix des représentants. Puisque $a/1 = c/1 \Longleftrightarrow (a - c)\, u = 0$ pour un certain $u \in M$, l'homomorphisme des anneaux $a \longmapsto a/1$ n'est un monomorphisme (un plongement) que pour l'anneau intègre K et son sous-monoïde M ne contenant pas zéro).

6. Appliquer la construction de $Q_M(K)$ à l'anneau $K = \mathbb{Z}$ et au monoïde $M = \mathbb{Z} \setminus p\mathbb{Z}$ formé de tous les entiers qui ne sont pas divisibles par un nombre premier fixe p. Montrer qu'on peut identifier $Q_M(\mathbb{Z})$ à l'ensemble de tous les nombres rationnels a/b, avec b non divisible par p.

CHAPITRE 6

ZÉROS DES POLYNÔMES

Nous abordons maintenant un sujet pour lequel on étudiait jadis l'algèbre, à savoir les zéros (ou racines) des polynômes. Ce domaine a cessé d'être dominant en algèbre, mais personne ne s'avise de mettre en doute son importance. Le fait est qu'en mathématiques de nombreux problèmes se ramènent finalement au calcul de certaines racines des polynômes concrets ou à la description qualitative de l'ensemble de ces racines. Nous ne pourrons traiter que les plus simples propriétés des racines, mais elles seront quand même suffisantes pour permettre d'apprécier, à juste titre, la place particulière occupée par le corps $\mathbb{C}$ des nombres complexes.

§ 1. Propriétés générales des racines

1. Racines et facteurs linéaires.— Soit A un anneau commutatif unitaire contenu dans un anneau intègre K.

Définition. — *On dit qu'un élément* $c \in K$ *est un zéro* (*une racine*) *du polynôme* $f \in A[X]$ *si* $f(c) = 0$. *On dit aussi que c'est une racine de l'équation* $f(x) = 0$.

La nécessité de considérer les anneaux dont A est une partie propre, sera compréhensible si l'on se rappelle que le polynôme $f(X) = X^2 + 1$ sur $\mathbb{R}$ n'a pas de zéros dans $\mathbb{R}$, alors que $f(i) = 0$, $i \in \mathbb{C} = \mathbb{R}[i]$. Commençons pourtant par considérer le cas où $K = A$.

Théorème 1 (théorème de Bézout). — *Un élément* $c \in A$ *est zéro d'un polynôme* $f \in A[X]$ *si, et seulement si,* $X - c$ *divise* f *dans l'anneau* $A[X]$.

Démonstration.— Ce théorème présente un cas particulier d'une assertion plus générale que nous aurions pu démontrer depuis longtemps. A savoir, l'algorithme de division euclidienne (voir chap. 5, § 2, théorème 5) dit que $f(X) = (X - c)\, q(X) + r(X)$, où $\deg r(X) < \deg (X - c) = 1$. Par conséquent, $r(X)$ est une cons-

15—0877

tante. La substitution de c à X (c'est-à-dire l'utilisation de l'application Π_c du théorème 2, § 2, chap. 5) donne $f(c) = r$, si bien qu'on a toujours

$$f(X) = (X - c)\, q(X) + f(c). \tag{1}$$

En particulier, $f(c) = 0 \Leftrightarrow f(X) = (X - c)\, q(X)$. ■

Effectuant la division d'un polynôme $f(X)$ à coefficients dans un anneau intègre A par un polynôme linéaire $X - c$, il est commode d'utiliser un schéma, dit *schéma de Horner*, qui est plus simple que l'algorithme général de division euclidienne. A savoir, soit

$$f(X) = a_0X^n + a_1X^{n-1} + \ldots + a_n, \quad a_i \in A.$$

D'après la formule (1), on a

$$q(X) = b_0X^{n-1} + b_1X^{n-2} + \ldots + b_{n-1}, \quad b_j \in A.$$

En comparant dans (1) les coefficients des mêmes puissances de X (en commençant par les coefficients dominants) on obtient, après une simple transformation,

$b_0 = a_0$	$\ldots$	$b_k = b_{k-1}c + a_k$	$\ldots$	$b_{n-1} = b_{n-2}c + a_{n-1}$	$f(c) = b_{n-1}c + a_n$	, (2)

si bien qu'on calcule à la fois la valeur de f pour $X = c$. Les formules récurrentes (2) qui traduisent justement le « schéma de Horner » s'avèrent bien commodes pour le calcul.

Etant donné le théorème 1, il est logique d'introduire une définition plus générale suivante :

DÉFINITION.— *On dit qu'un élément $c \in A$ est un zéro d'ordre de multiplicité k (une racine d'ordre de multiplicité k) du polynôme $f \in A[X]$ si f est divisible par $(X - c)^k$, mais n'est pas divisible par $(X - c)^{k+1}$. Si $k = 1$, on dit que c'est un zéro (une racine) simple (respectivement, si $k = 2$ et 3 on dit que c'est un zéro double et triple).*

Ainsi, $c \in A$ est un zéro d'ordre de multiplicité k du polynôme $f \in A[X]$ si, et seulement si, $f(X) = (X - c)^k g(X)$, où P.G.C.D.$(X - c, g(X)) = 1$. Eu égard à la formule 1, la dernière condition s'exprime aussi par l'inégalité $g(c) \neq 0$. En tenant compte du théorème 1, § 2, chapitre 5, remarquons que $\deg f = k + \deg g$, d'où $k \leqslant \deg f$. On peut énoncer le théorème important suivant :

THÉORÈME 2. — *Soient A un anneau intègre, $f \neq 0$ un polynôme de $A[X]$ et $c_1, \ldots, c_r$ ses zéros dans A d'ordres de multiplicité respectifs $k_1, \ldots, k_r$. Alors, on a*

$$f(X) = (X - c_1)^{k_1} \ldots (X - c_r)^{k_r}\, g(X),$$

où $g(X) \in A[X]$, $g(c_i) \neq 0$, $i = 1, \ldots, r$.

En particulier, le nombre de zéros distincts du polynôme $f \in A[X]$, considérés avec leurs ordres de multiplicité, est au plus égal au degré du polynôme:

$$k_1 + k_2 + \ldots + k_r \leqslant \deg f. \tag{3}$$

DÉMONSTRATION. — Il suffit de passer au corps des quotients $Q(A)$ (si l'anneau A n'était pas, dès le début, un corps) et d'utiliser l'unicité de la décomposition en facteurs premiers (dans le cas considéré, de la décomposition en $X - c_1, \ldots, X - c_r$) dans l'anneau $Q(A)[X]$ (voir résultats obtenus au chap. 5, §§ 3 et 4). Cependant, pour l'instant, on n'a pas besoin d'avoir recours à une arme si puissante. Raisonnons directement.

Puisque $\deg f = (k_1 + \ldots + k_r) + \deg g$, l'inégalité (3) est une conséquence de la divisibilité de f par $(X - c_1)^{k_1} \ldots (X - c_r)^{k_r}$ que nous établirons par récurrence sur r. Pour $r = 1$, il n'y a rien à démontrer. Supposons que nous sachions déjà que

$$f(X) = (X - c_1)^{k_1} \ldots (X - c_{r-1})^{k_{r-1}} h(X).$$

Etant donné que $c_r - c_1 \neq 0, \ldots, c_r - c_{r-1} \neq 0$ et A est un anneau intègre, l'élément c_r n'est pas un zéro du polynôme $(X - c_1)^{k_1} \ldots (X - c_{r-1})^{k_{r-1}}$. Or, c_r est un zéro de multiplicité k_r du polynôme f, c'est-à-dire $f(X) = (X - c_r)^{k_r} u(X)$. Donc, $h(c_r) = 0$. Respectivement, $h(X) = (X - c_r)^s v(X)$, $s \leqslant k_r$. On a

$$(X - c_r)^{k_r} u(X) = f(X) =$$
$$= (X - c_1)^{k_1} \ldots (X - c_{r-1})^{k_{r-1}} (X - c_r)^s v(X).$$

En utilisant la loi de simplification dans l'anneau intègre $A[X]$, on conclut que $s = k_r$. ■

Si l'anneau A n'est pas supposé intègre, le théorème 2 cesse d'être vrai comme le montre l'exemple du polynôme $f(X) = X^3$ sur l'anneau Z_8: $f(0) = f(2) = f(4) = f(6) = 0$. La décomposition de f en facteurs premiers dans $Z_8[X]$ n'est pas unique non plus: $f = X^3 = X(X - 4)^2 = (X - 2)(X^2 + 2X + 4) = (X - 6) \times$ $\times (X^2 - 2X + 4)$.

Du théorème 2 résulte le corollaire suivant:

COROLLAIRE.— *Si deux polynômes $f, g \in A[X]$ de degré $\leqslant n$ prennent les mêmes valeurs lors de la substitution de $n + 1$ éléments distincts de l'anneau intègre A, ils sont identiques: $f = g$.*

DÉMONSTRATION. — Posons $h = f - g$, de sorte que $\deg h \leqslant n$. Par hypothèse, $h(c_1) = \ldots = h(c_{n+1}) = 0$ pour des éléments deux à deux distincts $c_1, \ldots, c_{n+1} \in A$, c'est-à-dire le polynôme h de degré $\leqslant n$ a au moins $n + 1$ zéros. Cela nous conduit à une contradiction avec l'inégalité (3), contradiction qui ne peut être éliminée qu'en reconnaissant que $h = 0$. ■

2. Fonctions polynomiales.— Le corollaire du théorème 2 permet de résoudre le problème que nous avons abordé plus haut (voir chap. 5, § 2, nº 1), à savoir, le problème de la relation qui existe entre les notions de polynôme adoptées en théorie des fonctions et en algèbre. A tout polynôme $f \in A\,[X]$ nous faisons correspondre une fonction

$$\tilde{f}\colon a \mapsto f\,(a), \quad \forall a \in A.$$

L'ensemble de telles fonctions forme un anneau A_{pol} des *fonctions polynomiales* (on dit encore des *fonctions polynômes* ou des *fonctions rationnelles entières*), qui est un sous-anneau de l'anneau des fonctions $A^A = \{A \to A\}$ avec l'addition et la multiplication (voir chap. 4, § 4, nº 1, exemple 3 et chap. 5, § 2, théorème 2). En opérant tout à fait de la même manière, on introduit les fonctions polynomiales à plusieurs variables indépendantes.

Comme cela a été dit plus haut, le polynôme non nul $X^2 + X \in \mathbb{F}_2\,[X]$ définit une fonction nulle. En général, si $f\,(X) = (X^p - X)\,g\,(X)$ est un polynôme sur un corps fini à p éléments, $\tilde{f}$ est une fonction nulle, parce que $x^p - x = x\,(x^{p-1} - 1) = 0$ pour tous les $x \in \mathbb{F}_p$. C'est seulement dans le cas où $\deg f \leqslant p - 1$, que le polynôme $f \in \mathbb{F}_p\,[X]$ est défini par sa fonction $\tilde{f}$. Un polynôme arbitraire $f \in \mathbb{F}_p\,[X]$ peut être remplacé par un *polynôme réduit* f^* de degré $\leqslant p - 1$, univoquement déterminé, si l'on prend pour f^* le reste de la division de f par $X^p - X$. Il est alors évident que $\tilde{f} = \tilde{f}^*$.

Dans le cas des corps infinis ou des anneaux intègres infinis la situation est beaucoup plus simple.

THÉORÈME 3.— *Si A est un anneau intègre ayant un nombre infini d'éléments, l'application de l'anneau des polynômes $A\,[X]$ sur l'anneau des fonctions polynomiales A_{pol}, définie par la correspondance $f \mapsto \tilde{f}$, est un isomorphisme.*

Ce théorème n'est au fond qu'un autre énoncé du corollaire au théorème 2, puisqu'il s'agit seulement du fait qu'au polynôme $f \neq 0$ est associée une fonction non nulle $\tilde{f}$, c'est-à-dire que $f\,(a) \neq 0$ pour au moins un $a \in A$. Or, en réalité, f possède au plus n zéros dans A si $\deg f = n$. ■

En s'appuyant sur le théorème 3, on identifie l'anneau des polynômes sur un corps infini P à l'anneau des fonctions polynomiales notées $f\,(x)$ (avec un x minuscule), si bien qu'il ne reste qu'à répondre à la question : comment, à partir de $\tilde{f}$ (en réalité, à partir de quelques valeurs du polynôme f), rétablir sous une forme explicite le polynôme f lui-même.

Un énoncé correct du problème d'« interpolation » est le suivant. Soient $b_0, b_1, \ldots, b_n$ (respectivement, $c_0, c_1 \ldots, c_n$) $n + 1$ éléments quelconques (respectivement, distincts) d'un corps commutatif P. On demande de trouver un polynôme $f \in P\,[X]$ de degré $\leqslant n$, tel que $f\,(c_i) = b_i$, $i = 0, 1, \ldots, n$. Suivant le corollaire du théo-

rème 2, la solution de ce problème, si elle existe, est unique. Or, comme le montre la *formule d'interpolation de Lagrange*:

$$f(X)=\sum_{i=0}^{n} b_i \frac{(X-c_0)\ldots(X-c_{i-1})(X-c_{i+1})\ldots(X-c_n)}{(c_i-c_0)\ldots(c_i-c_{i-1})(c_i-c_{i+1})\ldots(c_i-c_n)} \qquad (4)$$

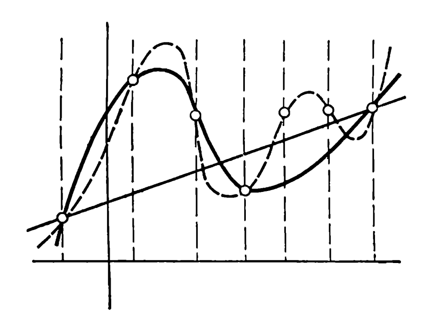
il existe toujours un polynôme f vérifiant les propriétés requises. D'ailleurs, l'existence et l'unicité de la solution découlent tout de suite du système linéaire

$$\begin{gathered} a_0c_0^n + a_1c_0^{n-1} + \ldots + a_n = b_0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_0c_n^n + a_1c_n^{n-1} + \ldots + a_n = b_n \end{gathered}$$

écrit pour les coefficients $a_0, a_1, \ldots, a_n$ du polynôme f cherché. Le déterminant de ce système, qui est un déterminant de Vandermonde, diffère de zéro, et les coefficients a_i se déterminent d'après les règles de Cramer. La formule (4) est bien commode parce qu'elle est simple et facile à retenir. Parfois, il est plus avantageux d'utiliser la *formule d'interpolation de Newton*:

$$\begin{aligned} f(X) = u_0 + u_1(X-c_0) + \ldots \\ \ldots + u_n(X-c_0)(X-c_1)\ldots(X-c_{n-1}), \qquad (5) \end{aligned}$$

dans laquelle les coefficients $u_0, u_1, \ldots, u_n$ sont déterminés par une substitution successive des valeurs $X = c_0$, $X = c_1, \ldots, X = c_n$. Les formules d'interpolation (4) et (5) sont pratiquement utilisées pour le calcul et la représentation graphique de la fonction $\varphi: \mathbb{R} \to \mathbb{R}$ donnée par un tableau de ses valeurs ou obtenue expérimentalement. Si l'on sait, à partir des considérations indirectes quelconques, que le comportement de la fonction φ sur un intervalle I de la droite réelle $\mathbb{R}$ est assez bon, on cherche à rapprocher φ sur I par une fonction aussi « lisse » que la fonction polynomiale.



En le faisant, on utilise comme « nœuds d'interpolation » des points $c_0, c_1, \ldots, c_n$, qui appartiennent à l'intervalle I, ces derniers

étant les (seuls) points, où sont connues les valeurs de φ : $\varphi(c_i) = b_i$. Les problèmes bien délicats du choix des nœuds d'interpolation et de l'élaboration des méthodes générales d'approximation des fonctions font l'objet de nombreux chapitres des mathématiques. Il y a lieu de noter que l'emploi de processus d'interpolation a joué un grand rôle dans le développement de la théorie des nombres transcendants (pour la définition des nombres algébriques et transcendants on se reportera au chap. 5, § 2), si bien que les intérêts de la théorie des fonctions, de la théorie des nombres et de l'algèbre s'y allient.

Notons en conclusion qu'à toute fraction rationnelle irréductible $f/g \in P(X)$ (voir chap. 5, § 4) et à toute extension $F \supset P$ à nombre infini d'éléments, est associée une *fonction rationnelle* $\widetilde{f/g}: F_{(f/g)} \to F$ à domaine de définition $F_{(f/g)}$ obtenu à partir de F en supprimant un nombre fini d'éléments, à savoir les zéros du polynôme g dans F. On peut démontrer que dans ces conditions l'application $f/g \mapsto \widetilde{f/g}$ est bijective. Nous n'aurons pas besoin de cette assertion. Intuitivement, elle est évidente. Malgré cette correspondance, il faut savoir distinguer très soigneusement les fonctions rationnelles des fractions rationnelles. La fonction rationnelle $x \mapsto 1/x$ n'est pas définie en $x = 0$, alors que la question, si la fraction rationnelle $1/X$ est définie ou non, ne se pose même pas.

3. Dérivations de l'anneau des polynômes.— La notion de polynôme adoptée en théorie des fonctions rend naturelle la définition suivante. Soit

$$f(X) = a_0X^n + a_1X^{n-1} + \ldots + a_{n-1}X + a_n$$

un polynôme de degré n sur un corps commutatif P. On appelle sa dérivée le polynôme

$$f'(X) = na_0X^{n-1} + (n-1)\,a_1X^{n-2} + \ldots + a_{n-1}. \qquad (6)$$

Si $P = \mathbb{R}$ est le corps des nombres réels et $\tilde{f}$ est la fonction polynomiale associée à f, la définition (6) coïncide avec la définition ordinaire de la dérivée en tant que limite:

$$\lim_{\Delta x \to 0} \frac{\tilde{f}(x+\Delta x) - \tilde{f}(x)}{\Delta x}.$$

Quant au cas où P est un corps commutatif arbitraire, il serait absurde de parler des propriétés quelconques de la continuité d'une fonction polynomiale (que doit-on entendre dans Z_p par suite convergente ?) et il faut donc se baser sur la définition formelle (6).

Pour la dérivation des polynômes sont valables les relations bien connues, établies en Analyse:

$$(\alpha f + \beta g)' = \alpha f' + \beta g', \quad \alpha, \beta \in P, \tag{7}$$

$$(fg)' = f'g + fg'. \tag{8}$$

La relation (7) découle directement de (6) et de la définition de la somme des polynômes. En se servant de (7) et de la définition du produit des polynômes, on peut ramener la vérification de (8) au cas où $f = X^k$, $g = X^l$:

$$(X^{k+l})' = (k+l)\, X^{k+l-1} = (kX^{k-1})\, X^l + X^k\,(lX^{l-1}) = \\ = (X^k)'\, X^l + X^k\,(X^l)'.$$

Une généralisation de (8) est donnée par la formule suivante qui est facile à démontrer par récurrence sur k:

$$(f_1 f_2 \ldots f_k)' = \sum_{i=1}^{k} f_1 \ldots f_{i-1} f_i' f_{i+1} \ldots f_k.$$

En particulier,

$$(f^k)' = kf^{k-1}f'. \tag{9}$$

Les relations (7), (8) écrites en termes d'application $\frac{d}{dX}: f \mapsto f'$ (on dit aussi que $\frac{d}{dX}$ est un *opérateur de dérivation*) incitent à introduire pour un anneau arbitraire K une application $\mathscr{D}: K \to K$, ayant les propriétés suivantes

$$\mathscr{D}(u + v) = \mathscr{D}u + \mathscr{D}v, \tag{7'}$$

$$\mathscr{D}(uv) = (\mathscr{D}u)\, v + u\, (\mathscr{D}v). \tag{8'}$$

De telles applications de l'anneau K dans lui-même, appelées *dérivations*, s'avèrent très utiles pour l'étude de K, alors que leur ensemble Der (K) est un être extrêmement intéressant qui introduit dans un vaste domaine des mathématiques (*groupes de Lie et algèbres de Lie*).

La relation (8′) est généralisée par la formule de Leibniz:

$$\mathscr{D}^m (uv) = \sum_{k=0}^{m} \binom{m}{k} \mathscr{D}^k u \mathscr{D}^{m-k} v, \tag{8''}$$

obtenue par récurrence sur $m \geqslant 1$ (l'application de $\mathscr{D}$ à (8″), l'utilisation de (8′) et la relation $\binom{m}{k-1} + \binom{m}{k} = \binom{m+1}{k}$ donnent (8″) pour $m + 1$).

Dans le cas où $K = P\,[X]$, les relations (7′), (8′), complétées de la règle

$$\mathscr{D}(\lambda f) = \lambda \mathscr{D} f, \quad \lambda \in P,$$

entraînent immédiatement que

$$\mathscr{D} f(X) = f'(X)\, \mathscr{D} X.$$

On voit donc que toute dérivation de l'anneau des polynômes $P[X]$ est définie par la donnée d'un seul polynôme $\mathscr{D} X$. Pour $\mathscr{D} X = 1$, on obtient l'opérateur de dérivation ordinaire $\frac{d}{dX}$.

4. Facteurs multiples.— En appliquant successivement m fois l'opérateur $\frac{d}{dX}$ au polynôme $f(X)$, on obtient le résultat désigné généralement par le symbole $f^{(m)}(X)$. Il est évident que

$$f(X) = a_0 X^n + a_1 X^{n-1} + \ldots + a_n \Rightarrow f^{(n)}(X) = n!\, a_0, \quad f^{(n+1)}(X) = 0.$$

Si P est un corps commutatif de caractéristique nulle, on a

$$\deg f' = \deg f - 1.$$

Cependant, il n'en est plus de même pour les corps de caractéristique finie p, parce que

$$(X^{kp})' = kpX^{kp-1} = 0.$$

Pourtant, l'étude de la dérivée permet de tirer un certain profit même dans le cas général. En divisant un polynôme quelconque $f \in P[X]$ par $(X - c)^2$, $c \in F$, $F \supset P$, et en écrivant ensuite le reste (linéaire) sous la forme $(X - c)s + r$, où $s, r \in F$, on obtient les relations $f = (X - c)^2 t + (X - c)s + r$, $f' = (X - c)[2t + (X - c)t'] + s$. En y portant la valeur de $X = c$, il vient $r = f(c)$, $s = f'(c)$, c'est-à-dire

$$f(X) = (X - c)^2 f(X) + (X - c) f'(c) + f(c).$$

On peut donc énoncer le théorème suivant:

THÉORÈME 4. — *Soient P un corps commutatif arbitraire et F une extension quelconque de ce corps. Le polynôme $f \in P[X]$ possède un zéro multiple $c \in F$ si, et seulement si, $f(c) = f'(c) = 0$.* ■

EXEMPLE 1. — Dans tout corps commutatif de caractéristique p, le polynôme $X^n - 1$ n'a que des zéros simples si n n'est pas divisible par p. En effet, les zéros du polynôme dérivé nX^{n-1} ne peuvent pas être zéros de $X^n - 1$.

On suppose maintenant que P est un corps commutatif de caractéristique nulle et, sans restreindre la généralité, on peut entendre par P l'un des corps $\mathbb{Q}$, $\mathbb{R}$ ou $\mathbb{C}$. Le polynôme unitaire irréductible $p_i(X)$ figurant dans la décomposition

$$f(X) = \lambda p_1(X)^{k_1} \ldots p_i(X)^{k_i} \ldots p_r(x)^{k_r}, \quad \lambda \in P, \qquad (10)$$

du polynôme $f(X) \in P[X]$ s'appelle (par analogie avec la définition du zéro multiple) facteur de multiplicité k_i de f. Comme cela a été dit plus haut, la décomposition (10) est difficile à obtenir

dans la pratique. Décrivons succinctement une méthode basée sur la notion de dérivée et permettant de connaître si $f(X)$ possède des facteurs multiples sur un corps donné P (ou sur son extension).

THÉORÈME 5. — *Soit $p(X)$ un facteur irréductible de multiplicité k du polynôme $f \in P[X]$ ($k \geqslant 1$, $\deg p(X) \geqslant 1$). Alors, $p(X)$ est facteur de multiplicité ($k - 1$) du polynôme dérivé $f'(X)$. En particulier, pour $k = 1$, f' n'est pas divisible par $p(X)$.*

DÉMONSTRATION. — Conformément à l'énoncé, $f(X) = p(X)^k g(X)$, où P.G.C.D.$(p(X), g(X)) = 1$, c'est-à-dire $g(X)$ n'est pas divisible par $p(X)$. En appliquant les règles (8) et (9), on trouve

$$f'(X) = p(X)^{k-1}[kp'(X)g(X) + p(X)g'(X)].$$

Il suffit de montrer que le polynôme entre crochets n'est pas divisible par $p(X)$. S'il n'en était pas ainsi, le polynôme $kp'(X)g(X)$ serait divisible par $p(X)$, ce qui est pourtant impossible (voir chap. 5, § 3, corollaires des théorèmes 3 et 4), parce que $g(X)$ n'est pas divisible par $p(X)$, et $\deg kp'(X) < \deg p(X)$. ■

Il est à noter que la démonstration ci-dessus s'appuie aussi bien sur l'irréductibilité de $p(X)$ que sur la condition car $P = 0$.

COROLLAIRE 1. — *Pour un polynôme $f(X)$ à coefficients dans un corps P de caractéristique nulle, les deux conditions suivantes sont équivalentes :*

(i) *f possède dans une certaine extension $F \supset P$ du corps P un zéro de multiplicité k;*

(ii) $f^{(j)}(c) = 0$, $0 \leqslant j \leqslant k-1$, *mais* $f^{(k)}(c) \neq 0$.

Pour démontrer cette assertion, il faut appliquer k fois le théorème 5, ayant en vue le facteur linéaire $p(X) = X - c$ et, s'il y a lieu, en remplaçant dès le début, le corps P par son extension F contenant le zéro c. ■

COROLLAIRE 2. — *Si un polynôme $f \in P[X]$ de degré $\geqslant 1$ possède la décomposition* (10), *la décomposition du plus grand commun diviseur de f et de sa dérivée f' est de la forme :*

$$\text{P.G.C.D.}(f, f') = p_1(X)^{k_1-1} p_2(X)^{k_2-1} \ldots p_r(X)^{k_r-1} \qquad (11)$$

(*le* P.G.C.D. *peut toujours être considéré comme un polynôme unitaire*).

En effet, d'après le théorème 5, chacun des diviseurs premiers $p_i(X)$ du polynôme $f(X)$ à décomposition canonique (10) intervient dans la décomposition de $f'(X)$ avec l'exposant $k_i - 1$, c'est-à-dire

$$f'(X) = p_1(X)^{k_1-1} p_2(X)^{k_2-1} \ldots p_r(X)^{k_r-1} \cdot u(X),$$

où P.G.C.D. $(u, p_i) = 1$, $1 \leqslant i \leqslant r$ (on suppose que $p_i(X)^0 = 1$). Aussi, en partant du critère de divisibilité déjà rencontré (voir

chap. 5, § 3, n° 2), peut-on conclure que le P.G.C.D. (f, f') se calcule par la formule (11). ■

En se servant de l'expression (11) pour le P.G.C.D. (f, f'), on obtient un moyen permettant de s'affranchir des facteurs multiples figurant dans la décomposition de $f(X)$. A savoir, le polynôme

$$g(X) = \frac{f(X)}{\text{P.G.C.D. } (f, f')} = p_1(X)\, p_2(X) \ldots p_r(X)$$

contient les mêmes diviseurs premiers que $f(X)$, ces derniers intervenant avec la multiplicité 1. Il importe de noter qu'on peut trouver le polynôme $g(X)$ sans connaître au fait les décompositions de f et f', uniquement à l'aide de l'algorithme d'Euclide.

EXEMPLE 2. — Le polynôme $f(X) = X^5 - 3X^4 + 2X^3 + 2X^2 - 3X + 1$ et son polynôme dérivé $f'(X) = 5X^4 - 12X^3 + 6X^2 + 4X - 3$ possèdent comme P.G.C.D. le polynôme unitaire $X^3 - 3X^2 + 3X - 1 = (X-1)^3$. Le polynôme « libre de carrés » $g(X) = f(X)/(X-1)^3 = X^2 - 1 = (X-1) \times \times (X+1)$ possède deux zéros ± 1. Ainsi, $f(X) = (X-1)^4 (X+1)$ a un zéro $+1$ de ultiplicité 4 et un zéro simple -1.

5. Formules de Viète.— Nous avons déjà parlé, à l'occasion de la théorie des systèmes d'équations linéaires, de l'influence favorable qu'a eue sur le développement de cette théorie un bon système de notations qui a amené en particulier aux déterminants. Ce mérite revient aux mathématiciens du XVIIIe et du début du XIXe siècles. Or, beaucoup plus tôt, à l'époque où l'algèbre était encore identifiée avec « l'analyse des équations », un perfectionnement décisif des notations algébriques, réalisé par Viète et Descartes, a grandement favorisé le développement de la théorie des polynômes et des équations algébriques. Des cas particuliers d'équations à coefficients numériques qui voilaient les lois générales, on a passé résolument aux équations à coefficients littéraux. Il n'est pas rare qu'un nouveau procédé d'écriture fait naître de nouveaux résultats. Dans les travaux de Descartes, cela a été couronné par une application révolutionnaire de l'algèbre à la géométrie. Nous passerons en revue un résultat plus modeste, acquis par son prédécesseur Viète.

Supposons qu'un polynôme unitaire $f \in P[X]$ de degré n possède dans le corps commutatif P ou dans une extension de ce corps n zéros $c_1, c_2, \ldots, c_n$ dont quelques-uns sont peut-être identiques. Alors, en vertu du théorème 2, on a la décomposition

$$f(X) = (X - c_1)(X - c_2) \ldots (X - c_n).$$

Ecrivons $f(X)$ sous la forme usuelle, suivant les puissances de X:

$$f(X) = X^n + a_1 X^{n-1} + \ldots + a_k X^{n-k} + \ldots + a_n,$$

à cet effet, multiplions tous les binômes $X - c_i$ et réduisons les termes semblables. Les coefficients $a_1, \ldots, a_n$ s'expriment alors

en fonction des zéros $c_1, \ldots, c_n$:

$$\begin{aligned} a_1 &= -(c_1 + c_2 + \ldots + c_n), \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ a_k &= (-1)^k \sum_{i_1<i_2<\ldots<i_k} c_{i_1}c_{i_2}\ldots c_{i_k}, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ a_n &= (-1)^n c_1c_2\ldots c_n. \end{aligned} \tag{12}$$

Les formules (12) sont connues sous le nom de *formules de Viète*.

Si le polynôme f n'était pas unitaire, c'est-à-dire si son coefficient dominant était $a_0 \neq 1$, les formules analogues à (12) donneraient les expressions pour les rapports a_i/a_0.

Les formules de Viète qui établissent une liaison explicite entre les zéros et les coefficients d'un polynôme arbitraire (unitaire), sont remarquables par ce que leurs seconds membres ne changent pas pour toute permutation des zéros $c_1, \ldots, c_n$. Cela nous permet d'introduire la notion de *fonction symétrique*, tout comme à l'occasion des déterminants il s'est trouvé commode de considérer des fonctions symétriques gauches générales. En vertu de la définition donnée dans le corollaire au théorème 3' du chapitre 5, § 2, nº 2, un élément π du groupe symétrique S_n opère sur la fonction $\tilde{f}(x_1, \ldots, x_n)$ à n arguments, suivant la règle:

$$(\widetilde{\pi f})(x_1, \ldots, x_n) = \tilde{f}(x_{\pi^{-1}(1)}, \ldots, x_{\pi^{-1}(n)}).$$

On dit que la fonction $\tilde{f}$ est symétrique si $\pi\tilde{f} = \tilde{f}$ pour tout $\pi \in S_n$. A titre d'exemple des fonctions symétriques, on peut indiquer les *fonctions* s_k dites *symétriques élémentaires*:

$$s_k(x_1, \ldots, x_n) = \sum_{1\leqslant i_1<i_2<\ldots<i_k\leqslant n} x_{i_1}x_{i_2}\ldots x_{i_k}. \tag{13}$$

Elles permettent de mettre les formules (12) sous la forme

$$a_k = (-1)^k s_k(c_1, \ldots, c_n),\ k = 1, 2, \ldots, n, \tag{12'}$$

si bien que le coefficient a_k du polynôme f est, au signe près, la valeur de la fonction s_k sur l'ensemble des zéros du polynôme f. Portons notre attention sur le fait que, par définition, $a_k \in P$, bien que dans le cas général les zéros $c_1, \ldots, c_n$ appartiennent à une certaine extension $F \supset P$. Quant à la question relative à l'existence de F, nous la laissons pour l'instant. Mais parfois, la décomposition d'un polynôme en facteurs linéaires est une conséquence directe des propriétés du corps P.

EXEMPLE. — Considérons le polynôme $X^{p-1} - 1$ sur le corps commutatif fini $\mathbb{F}_p$ (voir chap. 4, § 4, nº 6). Nous savons que $x^{p-1} = 1$ pour tout $x \in \mathbb{F}_p^*$, c'est-à-dire que tous les éléments non nuls sont zéros du polynôme $X^{p-1} - 1$.

On a donc la décomposition

$$X^{p-1} - 1 = (X - 1)(X - 2) \dots (X - (p - 1)). \qquad (14)$$

On suppose que nous nous sommes familiarisés avec le corps $\mathbb{F}_p$ au point de pouvoir distinguer la dualité des nombres $1, 2, \dots, p - 1$ qui sont des éléments ordinaires de $\mathbb{Z}$ d'une part, et des éléments du corps $\mathbb{F}_p \cong \mathbb{Z}/p\,\mathbb{Z}$ (représentants des classes résiduelles $\{i\}_p$) d'autre part. Si $p > 2$, on déduit de (12′) que

$$s_k(1, 2, \dots, p - 1) \equiv 0 \pmod{p}, \quad k = 1, 2, \dots, p - 2,$$

$$s_{p-1}(1, 2, \dots, p - 1) \equiv -1 \pmod{p}.$$

La dernière relation mise sous la forme

$$(p - 1)! + 1 \equiv 0 \pmod{p} \qquad (15)$$

est connue sous le nom de *théorème de Wilson*. Elle exprime au fait une condition (un critère) nécessaire et suffisante pour qu'un entier p soit premier. En effet, nous venons de démontrer que la relation (15) est vérifiée pour les p premiers. D'autre part, $p = p_1p_2 \Rightarrow (p - 1)! = p_1t \Rightarrow (p - 1)! + 1 \not\equiv 0 \pmod{p_1} \Rightarrow$ $\Rightarrow (p - 1)! + 1 \not\equiv 0 \pmod{p}$. ■

EXERCICES

1. L'anneau des fonctions polynomiales sur un corps à p éléments est-il intègre ?

2. Soient P un corps commutatif infini et f un polynôme non nul de $P[X_1, \dots, X_n]$. En s'appuyant sur le théorème 3 et en raisonnant par récurrence sur n, démontrer l'existence de $a_1, \dots, a_n \in P$ pour lesquels $f(a_1, \dots, a_n) \neq 0$. Il en résulte un isomorphisme entre $P[X_1, \dots, X_n]$ et l'anneau des fonctions polynomiales à n indéterminées sur P.

3. Un polynôme non nul $f \in Z_p[X_1, \dots, X_n]$ de degré $<p$ par rapport à chaque indéterminée, vérifie la propriété énoncée dans l'exercice 2: $f(a_1, \dots, a_n) \neq 0$ pour certains $a_1, \dots, a_n \in Z_p$. Montrer que tout polynôme $f \in Z_p[X_1, \dots, X_n]$ peut s'écrire sous la forme

$$f(X_1, \dots, X_n) = \sum_{i=1}^{n} g_i(X_1, \dots, X_n)(X_i^p - X_i) + f^*(X_1, \dots X_n),$$

où f^* est un polynôme **réduit** ($\deg_{X_i} f^* \leqslant p - 1$, $i = 1, 2, \dots, n$) de degré $\deg f^* \leqslant \deg f$. Tirer une conclusion justifiée que l'application $f \mapsto \tilde{f} = \tilde{f}^*$ est un épimorphisme de l'anneau $Z_p[X_1, \dots, X_n]$ sur l'anneau des fonctions polynomiales à n indéterminées sur Z_p, avec noyau $L = \sum_{i=1}^{n} (X_i^p - X_i) \times$ $\times Z_p[X_1, \dots, X_n]$.

4. THÉORÈME (Chevalley).— *Soit $f(X_1, \dots, X_n)$ un polynôme homogène (une forme) sur Z_p de degré $r < n$. Alors, l'équation $f(x_1, \dots, x_n) = 0$ possède au moins une solution non triviale.* (I n d i c a t i o n. Puisque f est une forme, il est évident que $f(0, \dots, 0) = 0$. En raisonnant par l'absurde, supposer que $(a_1, \dots, a_n) \neq (0, \dots, 0) \Rightarrow f(a_1, \dots, a_n) \neq 0$. En partant de l'exercice 3 et du théorème de Fermat, en déduire que le polynôme réduit, associé à $g(X_1, \dots, X_n) = 1 - f(X_1, \dots, X_n)^{p-1}$, sera de la forme $g^*(X_1, \dots, X_n) =$ $= (1 - X_1^{p-1}) \dots (1 - X_n^{p-1})$. Or,

$$\deg g = (p - 1) \deg f = (p - 1)\, r < (p - 1)\, n = \deg g^*.$$

La contradiction obtenue démontre le théorème.)

En modifiant légèrement le raisonnement (ayant calculé par deux procédés la somme $\sum_{x_1, \ldots, x_n \in Z_p} g(x_1, \ldots, x_n)$) démontrer que le nombre total de solutions est toujours divisible par p.

5. Soit $f(x_1, \ldots, x_n)$ une forme quadratique à coefficients entiers. Le théorème de Chevalley (voir exercice 4), énoncé en termes de la théorie des congruences, affirme que pour $n \geqslant 3$, la congruence

$$f(x_1, \ldots, x_n) \equiv 0 \pmod{p}$$

possède une solution non nulle. Vérifier que toutes les solutions de la congruence $x^2 - 2y^2 \equiv 0 \pmod{5}$ sont triviales, et que la condition $r < n$ est donc nécessaire.

6. Montrer que P.G.C.D. $(f', f) = 1$ si car $P = 0$, f est un polynôme irréductible sur un corps commutatif P et f' sa dérivée.

7. Démontrer que $f' = 0 \Rightarrow f = \text{const}$ pour un polynôme $f(X)$ sur un corps de caractéristique nulle, et $f' = 0 \Rightarrow f(X) = g(X^p)$ pour un polynôme $f(X)$ sur un corps de caractéristique $p > 0$ (g est un autre polynôme).

8. Au nº 3 nous avons vu que toute dérivation de l'anneau des polynômes $P[X]$ est de la forme $T_u: f \mapsto uf'$, $u \in P[X]$. Démontrer les assertions suivantes:

(i) l'ensemble des constantes (ce qui s'annule lors de la dérivation) est un sous-anneau de $P[X]$;

(ii) le produit $T_u T_v$ n'est pas en général une dérivation, mais si car $P = p > 0$, la puissance $(T_u)^p$ est une dérivation;

(iii) le commutateur $[T_u, T_v] = T_u T_v - T_v T_u$ est toujours une dérivation de la forme T_w, où $w = uv' - u'v$.

9. Dans le cas d'un anneau $P[X_1, \ldots, X_n]$ des polynômes à n indéterminées, il est naturel d'introduire un *opérateur de dérivation partielle par rapport à la k-ième indéterminée*:

$$\frac{\partial}{\partial X_k}: X_1^{i_1} \ldots X_k^{i_k} \ldots X_n^{i_n} \mapsto i_k X_1^{i_1} \ldots X_k^{i_k - 1} \ldots X_n^{i_n}.$$

(i) Montrer que l'ensemble des constantes pour $\frac{\partial}{\partial X_k}$ est l'anneau $P[X_1, \ldots, \hat{X}_k, \ldots, X_n]$ des polynômes à $n - 1$ indéterminées.

(ii) Soit $f(X_1, \ldots, X_n)$ une forme (un polynôme homogène) de degré m. S'assurer que l'*identité d'Euler*

$$\sum_{k=1}^{n} X_k \frac{\partial f}{\partial X_k} = m \cdot f(X_1, \ldots, X_n)$$

est vraie. Réciproquement, si car $P = 0$, l'identité d'Euler n'est vérifiée que par les formes de degré $m = 1, 2, 3, \ldots$

10. Montrer que l'absence de facteurs linéaires dans la décomposition du polynôme $X^n + a_1 X^{n-1} + \ldots + a_n \in Z_2[X]$ est une condition nécessaire et suffisante pour que soit vérifiée la relation $a_n (1 + \sum a_i) \neq 0$. Pour $n \leqslant 3$, les seuls polynômes irréductibles sur Z_2 sont les suivants: X, $X + 1$, $X^2 + X + 1$, $X^3 + X + 1$, $X^3 + X^2 + 1$. Ecrire tous les polynômes irréductibles sur Z_2 pour $n = 4$ et 5 (ils seront au nombre de 3 et 6 respectivement).

11. En partant de la congruence

$$X^5 - X - 1 \equiv (X^3 + X^2 + 1)(X^2 + X + 1) \pmod{2}$$

établir que le polynôme $X^5 - X - 1$ est irréductible sur $\mathbb{Q}$. (Indication. Appliquer le corollaire du lemme de Gauss (chap. 5, § 3), utiliser l'exercice précédent et tenir compte du fait que l'anneau $Z_2[X]$ est factoriel.)

Procédant de façon analogue, démontrer que le polynôme $X^5 - X^2 - 1$ est irréductible sur $\mathbb{Q}$ après avoir constater qu'il est irréductible modulo 2.

§ 2. Polynômes symétriques

1. Anneau des polynômes symétriques.— Suivant la définition des fonctions symétriques qui a été donnée à la fin du paragraphe précédent, nous introduirons une notion analogue dans l'anneau $A[X_1, \ldots, X_n]$ des polynômes sur un anneau intègre A. Il semble à première vue qu'étant étendu aux polynômes et aux fonctions à plusieurs indéterminées, le théorème 3 du § 1 rend inutile un tel report. Or, il faut ne pas perdre de vue que dans ce théorème l'anneau intègre A de coefficients est infini, alors que nous voulons avoir une construction universelle.

Ainsi, en revenant au corollaire du théorème 3' (voir chap. 5, § 2, n° 2), nous faisons correspondre à chaque permutation $\pi \in S_n$ un automorphisme $\tilde{\pi}: A[X_1, \ldots, X_n] \to A[X_1, \ldots, X_n]$ qui transforme un polynôme arbitraire $f \in A[X_1, \ldots, X_n]$ en polynôme πf:

$$(\pi f)(X_1, \ldots, X_n) = f(X_{\pi^{-1}(1)}, \ldots, X_{\pi^{-1}(n)}).$$

On dit que le polynôme f est *symétrique* si $\pi f = f$ pour tous les $\pi \in S_n$. De même que dans le cas des fonctions, on introduit les polynômes symétriques élémentaires s_k:

$$s_k(X_1, \ldots, X_n) = \sum_{1 \leqslant i_1 < i_2 < \ldots < i_k \leqslant n} X_{i_1} X_{i_2} \ldots X_{i_k}, \qquad k = 1, 2, \ldots, n. \tag{1}$$

Il faudrait, en général, considérér le polynôme

$$f(Y) = (Y - X_1)(Y - X_2) \ldots (Y - X_n) = Y^n - s_1 Y^{n-1} + s_2 Y^{n-2} + \ldots + (-1)^n s_n \tag{2}$$

sur $A[X_1, \ldots, X_n]$ à une nouvelle indéterminée Y et remarquer que s_k est un polynôme symétrique, car le premier membre de l'identité (2) reste inchangé pour toute permutation des facteurs linéaires $Y - X_1, \ldots, Y - X_n$.

Portons notre attention sur le fait qu'en substituant zéro à X_n dans les deux membres de l'identité (2), on obtient $(Y - X_1) \ldots \ldots (Y - X_{n-1}) Y = Y^n - (s_1)_0 Y^{n-1} + \ldots + (-1)^{n-1} (s_{n-1})_0 Y$, où $(s_k)_0$ est le résultat de la substitution $X_n = 0$ dans s_k. En simplifiant les deux membres par Y (en vertu du théorème 3 du chap. 4, § 4 et du théorème 1 du chap. 5, § 2, appliqués à $A[X_1, \ldots, X_n,$

$Y]$), on est conduit à l'identité

$$(Y - X_1)(Y - X_2) \dots (Y - X_{n-1}) =$$
$$= Y^{n-1} - (s_1)_0 Y^{n-2} + \dots + (-1)^{n-1} (s_{n-1})_0. \quad (3)$$

En comparant (2) et (3) on arrive à la conclusion que $(s_1)_0, \dots, (s_{n-1})_0$ *sont des polynômes symétriques élémentaires à* $n - 1$ *indéterminées* $X_1, \dots, X_{n-1}$.

Etant donné que π est un automorphisme de l'anneau $A [X_1, \dots, X_n]$, toute combinaison linéaire de polynômes symétriques, ainsi que tous leurs produits seront de nouveau des polynômes symétriques. Cela signifie que *l'ensemble de tous les polynômes symétriques forme un anneau qui est un sous-anneau de l'anneau* $A [X_1, \dots, X_n]$. Proposons-nous maintenant d'étudier tout d'abord comment est construit ce sous-anneau.

2. Théorème fondamental sur les polynômes symétriques.— Il se trouve que le procédé le plus général, permettant d'obtenir les polynômes symétriques, est le suivant. Il faut prendre un polynôme arbitraire $g \in A [Y_1, \dots, Y_n]$ et y substituer respectivement $s_1, \dots, s_n$ à $Y_1, \dots, Y_n$. Le polynôme

$$f(X_1, \dots, X_n) = g(s_1(X_1, \dots, X_n), \dots, s_n(X_1, \dots, X_n))$$

qui en résultera sera, certes, symétrique.

Notons encore que le monôme $Y_1^{i_1} \dots Y_n^{i_n}$ faisant partie de g, se transforme lors de la substitution $Y_k = s_k(X_1, \dots, X_n)$ en un polynôme homogène à indéterminées $X_1, \dots, X_n$ de degré $i_1 + 2i_2 + \dots + ni_n$, car $\deg s_k = k$. La somme $i_1 + 2i_2 + \dots + ni_n$ est généralement appelée *poids du monôme* $Y_1^{i_1} \dots Y_n^{i_n}$. Il est naturel d'appeler *poids du polynôme* $g(Y_1, \dots, Y_n)$ le plus grand des poids des monômes figurant dans g.

L'assertion fondamentale relative aux polynômes symétriques est exprimée par le théorème suivant:

THÉORÈME 1. — *Soit* $f \in A [X_1, \dots, X_n]$ *un polynôme symétrique de degré total* m *sur un anneau intègre* A. *Alors, il existe un polynôme* $g \in A [Y_1, \dots, Y_n]$ *de poids* m, *et un seul, tel que l'on ait*:

$$f(X_1, \dots, X_n) = g(s_1, \dots, s_n).$$

La démonstration se compose de deux parties.

I. DÉMONSTRATION DE L'EXISTENCE DU POLYNÔME g. — Raisonnons par récurrence sur deux paramètres n et m (voir chap. 1, § 7). Pour $n = 1$, le théorème est évident parce que $s_1 = X_1$ et $f(X_1) = f(s_1)$. En supposant que l'assertion sur l'existence de g soit démontrée pour les polynômes à $\leqslant n - 1$ indéterminées, raisonnons, dans le cas de n indéterminées, par récurrence sur $m = \deg f$. Puisque, pour $m = 0$,

il n'y a rien à démontrer, supposons $m > 0$ et considérons comme établie l'existence de g pour tout polynôme de degré $<m$.

Soit maintenant $f(X_1, \ldots, X_n)$ un polynôme symétrique donné de degré m. En posant $X_n = 0$, on aura, par hypothèse de récurrence,

$$f(X_1, \ldots, X_{n-1}, 0) = g_1((s_1)_0, \ldots, (s_{n-1})_0),$$

où g_1 est un polynôme quelconque de $A[Y_1, \ldots, Y_{n-1}]$ de poids $\leqslant m$ (le degré de f a pu s'abaisser lors de la substitution $X_n = 0$) et $(s_1)_0, \ldots, (s_{n-1})_0$ sont des polynômes symétriques élémentaires de $X_1, \ldots, X_{n-1}$ (voir (3)). Ceci étant, il est évident que $\deg g_1(s_1, \ldots \ldots, s_{n-1}) \leqslant m$. Ainsi donc, le polynôme

$$f_1(X_1, \ldots, X_n) = f(X_1, \ldots, X_n) - g_1(s_1, \ldots, s_{n-1}) \qquad (4)$$

a le degré total par rapport à $X_1, \ldots, X_n$ au plus égal à m, et est symétrique (en tant que différence de deux polynômes symétriques). De plus, $f_1(X_1, \ldots, X_{n-1}, 0) = 0$, d'où l'on déduit que X_n divise f_1: $f_1 = X_n \cdot f^\circ$. Or, il vient par suite de la symétrie que $f_1 = \pi^{-1}f_1 = X_{\pi(n)}(\pi^{-1}f^\circ)$, $\forall \pi \in S_n$, c'est-à-dire que f_1 contient comme facteurs $X_1, X_2, \ldots, X_n$, donc leur produit $s_n = X_1X_2 \ldots X_n$. Ainsi,

$$f_1(X_1, \ldots, X_n) = s_n \cdot f_2(X_1, \ldots, X_n), \qquad (5)$$

où f_2 est de nouveau un polynôme symétrique, cette fois de degré $\deg f_2 = \deg f_1 - n \leqslant m - n$. Par hypothèse de récurrence, il existe un polynôme $g_2(Y_1, \ldots, Y_n)$ de poids $\leqslant m - n$, pour lequel $f_2(X_1, \ldots, X_n) = g_2(s_1, \ldots, s_n)$. En tenant compte de (4) et (5), on obtient pour f l'expression

$$f(X_1, \ldots, X_n) = g_1(s_1, \ldots, s_{n-1}) + s_n g_2(s_1, \ldots, s_n).$$

Il s'ensuit que l'existence du polynôme $g = g_1(Y_1, \ldots, Y_n) + Y_n g_2(Y_1, \ldots, Y_n)$ de poids $\leqslant m$ est établie. Puisque $\deg f = m$, le poids du polynôme g ne peut pas être inférieur à m, il est donc égal exactement à m.

II. Démonstration de l'unicité. — S'il existait deux polynômes non identiques g_1, g_2 satisfaisant à la condition $f = g_1(s_1, \ldots \ldots, s_n) = g_2(s_1, \ldots, s_n)$, on aurait un polynôme $g(Y_1, \ldots \ldots, Y_n) = g_1 - g_2 \neq 0$ pour lequel $g(s_1, \ldots, s_n) = 0$. En d'autres termes, $s_1, \ldots, s_n$ seraient algébriquement dépendants sur A au sens de la définition donnée au chapitre 5, § 2, n° 2. Montrons (de nouveau par récurrence sur n) qu'il n'en est pas ainsi. En effet, en raisonnant par l'absurde, choisissons un polynôme $g(Y_1, \ldots \ldots, Y_n)$ de degré minimal, qui s'annule lors de la substitution $Y_k = s_k$. En considérant g en tant que polynôme en Y_n sur $A[Y_1, \ldots \ldots, Y_{n-1}]$, écrivons-le sous la forme

$$g(Y_1, \ldots, Y_n) =$$
$$= g_0(Y_1, \ldots, Y_{n-1}) + \ldots + g_k(Y_1, \ldots, Y_{n-1}) Y_n^k, \; k = \deg_n g.$$

Si $g_0 = 0$, alors $g = Y_n h$, où $h \in A[Y_1, \ldots, Y_n]$. Par hypothèse, $s_n h(s_1, \ldots, s_n) = 0$, et puisque l'anneau $A[X_1, \ldots, X_n]$ est intègre (chap. 5, § 2, théorème 1′), on en déduit l'égalité $h(s_1, \ldots$ $\ldots, s_n) = 0$. Pourtant, cela n'est pas possible, parce que $\deg h(Y_1, \ldots, Y_n) = \deg g(Y_1, \ldots, Y_n) - 1$. Donc, $g_0 \neq 0$. Considérons maintenant dans $A[X_1, \ldots, X_n]$ l'identité

$$g_0(s_1, \ldots, s_{n-1}) + \ldots + g_k(s_1, \ldots, s_{n-1})\, s_n^k = g(s_1, \ldots, s_n) = 0$$

et substituons 0 à X_n. Il en résulte que tous les termes sauf le premier s'annulent, et on aura

$$g_0((s_1)_0, \ldots, (s_{n-1})_0) = 0,$$

où $(s_1)_0, \ldots, (s_{n-1})_0$ sont les polynômes symétriques élémentaires à indéterminées $X_1, \ldots, X_{n-1}$ (voir (3)). Par hypothèse de récurrence, ils sont algébriquement indépendants sur A. En même temps, on a $g_0 \neq 0$. La contradiction obtenue démontre l'unicité du polynôme et donc le théorème 1. ∎

Remarquons que la démonstration de la première partie du théorème a été constructive; elle peut donc être utilisée pour la recherche pratique du polynôme g. En outre, les raisonnements faits au cours de la démonstration entraînent que les coefficients du polynôme cherché g sont contenus dans un sous-anneau de l'anneau A, engendré par les coefficients du polynôme donné f. En particulier, lorsque $A = \mathbb{Z}$, les coefficients des polynômes f et g seront des entiers.

COROLLAIRE. — *Soit* $f(X) = X^n + a_1 X^{n-1} + \ldots + a_{n-1}X + a_n$ *un polynôme unitaire de degré* n *à une indéterminée* X *sur un corps commutatif* P, *possédant* n *zéros* $c_1, \ldots, c_n$ *dans un certain corps plus grand* $F \supset P$. *Soit ensuite* $h(X_1, \ldots, X_n)$ *un polynôme symétrique quelconque de* $P[X_1, \ldots, X_n]$. *Alors, sa valeur* $h(c_1, \ldots$ $\ldots, c_n)$ *obtenue par suite de la substitution de* c_i *à* X_i, $i = 1, \ldots$ $\ldots, n$, *appartient au corps* P.

DÉMONSTRATION. — En effet, suivant le théorème fondamental sur les polynômes symétriques, il existe un polynôme $g(Y_1, \ldots$ $\ldots, Y_n) \in P[Y_1, \ldots, Y_n]$ tel que $h(X_1, \ldots, X_n) = g(s_1(X_1, \ldots$ $\ldots, X_n), \ldots, s_n(X_1, \ldots, X_n))$. C'est pourquoi $h(c_1, \ldots, c_n) = g(s_1(c_1, \ldots, c_n), \ldots, s_n(c_1, \ldots, c_n))$ et, puisque d'après les formules de Viète (12) du § 1, $s_k(c_1, \ldots, c_n) = (-1)^k a_k \in P$, on a aussi $g(-a_1, \ldots, (-1)^n a_n) \in P$. ∎

3. Méthode des coefficients indéterminés.— Il existe plusieurs démonstrations différentes du théorème fondamental sur les polynômes symétriques et donc, respectivement, plusieurs méthodes permettant d'exprimer un polynôme donné f par des polynômes symé-

triques élémentaires. Pour pouvoir décrire l'une des méthodes les plus employées, introduisons un nouveau type de polynômes symétriques. Pour fixer les idées, prenons pour A l'anneau $\mathbb{Z}$ ou le corps $\mathbb{R}$. Soit $v = X_1^{i_1} \ldots X_n^{i_n}$ un monôme quelconque. Convenons d'appeler v *monôme monotone*, si $i_1 \geqslant i_2 \geqslant \ldots \geqslant i_n$. Désignons par $S(v)$ la somme de tous les monômes différents dans la famille de $n!$ monômes de la forme πv, $\pi \in S_n$. Autrement dit

$$S(v) = \sum_{\pi \in S_n/H} \pi v,$$

où la condition $\pi \in S_n/H$ signifie que π parcourt l'ensemble des représentants des classes à gauche πH du groupe S_n suivant le sous-groupe $H = \{\tau \in S_n \mid \tau v = v\}$ (on peut vérifier aisément que le sous-ensemble H ainsi défini est réellement un sous-groupe). Par exemple

$$p_k(X_1, \ldots, X_n) = S(X_n^k) = X_1^k + X_2^k + \ldots + X_n^k, \quad k \geqslant 0, \qquad (6)$$

est une somme appelée *somme de puissances*. Il est évident qu'ici $H = S_{n-1}$. On a ensuite $S(X_1X_2 \ldots X_k) = s_k(X_1, \ldots, X_n)$ (avec quoi H coïncide-t-il ici ?). Il est clair que $S(v)$ est un polynôme homogène symétrique du même degré total que v. Puisque $S(v) = S(\sigma v)$, $\forall\, \sigma \in S_n$, il est logique de ne considérer que des sommes $S(v)$ ayant pour termes des monômes monotones v. D'après le sens, il est aussi clair que tout polynôme symétrique f sur A est une combinaison linéaire à coefficients dans A des polynômes de type $S(v)$:

$$f = \sum a_v S(v).$$

D'ordinaire, une telle écriture s'obtient instantanément (« au jugé »). Ainsi, le problème se ramène à exprimer $S(v)$ en fonction des polynômes symétriques élémentaires.

Convenons de disposer les monômes constitutifs de $S(v)$ *dans l'ordre lexicographique* (d'après le principe d'établissement d'un dictionnaire), c'est-à-dire de la manière suivante: le monôme $v = X_1^{i_1} X_2^{i_2} \ldots X_n^{i_n}$ précède le monôme $w = X_1^{j_1} X_2^{j_2} \ldots X_n^{j_n}$ (ou est supérieur à ce dernier, $v > w$) si et seulement si, la suite $i_1 - j_1$, $i_2 - j_2$, $\ldots$, $i_n - j_n$ est de la forme $0, \ldots, 0, t, \ldots$, où $t > 0$ (il peut arriver qu'à droite de t se placent aussi des différences négatives $i_l - j_l$). Il va de soi que ce principe lexicographique de disposition des termes s'applique non seulement à $S(v)$, mais à tout polynôme $f \in A[X_1, \ldots, X_n]$. Au sens lexicographique, le *premier* terme de la somme $S(v)$ à monôme monotone v sera v. Pour un monôme monotone $v = X_1^{i_1} X_2^{i_2} \ldots X_n^{i_n}$ nous sommes en droit de considérer le produit

$$g_v = s_1^{i_1 - i_2} s_2^{i_2 - i_3} \ldots s_{n-1}^{i_{n-1} - i_n} s_n^{i_n}, \quad s_i = s_i(X_1, \ldots, X_n), \qquad (7)$$

dans lequel le premier terme sera de nouveau

$$v = X_1^{i_1-i_2}(X_1X_2)^{i_2-i_3}\dots(X_1\dots X_{n-1})^{i_{n-1}-i_n}(X_1\dots X_n)^{i_n}$$

(pour obtenir le premier terme du produit, il faut prendre le premier terme de chacun des facteurs). On en déduit que le premier terme de la différence $S(v) - g_v$ sera inférieur à v. Donc,

$$S(v) - g_v = \sum n'_w S(w),$$

où $n'_w \in \mathbb{Z}$, et la sommation est étendue à l'ensemble des monômes monotones $w < v$. Les degrés totaux de v et de tous les w coïncident.

On peut maintenant esquisser la méthode suivante pour exprimer le polynôme $S(v)$ par des polynômes symétriques élémentaires. Soit $\deg v = m$. On prend toutes les partitions « monotones »

$$m = j_1 + j_2 + \dots + j_n,\ j_1 \geqslant j_2 \geqslant \dots \geqslant j_n \geqslant 0,$$

de l'entier m, telles que l'on ait $w = X_1^{j_1}X_2^{j_2}\dots X_n^{j_n} < v$. On considère l'ensemble M_v de tous les monômes w de ce type. Pour chaque $w \in M_v$ on compose le monôme g_w (voir (7)). Nous savons déjà que

$$S(v) = g_v + \sum_{w \in M_v} n_w g_w, \tag{8}$$

où n_w sont des entiers quelconques. Les coefficients indéterminés n_w (d'où l'appellation de « méthode des coefficients indéterminés ») sont calculés en substituant successivement, dans (8), à $X_1, \dots, X_n$, des entiers quelconques, généralement, les zéros et les unités. Les valeurs de g_v, g_w et $S(v)$ sont dans ce cas connues, si bien qu'on obtient pour n_w un système d'équations linéaires *a priori* compatible.

EXEMPLE. — Soit $v = X_1^3$, $S(v) = p_3(X_1, \dots, X_n)$, $n \geqslant 3$, $g_v = s_1^3$,

M_v	$X_1^2X_2$	$X_1X_2X_3$
g_w	s_1s_2	s_3

L'équation (8) est dans ce cas de la forme

$$p_3 = s_1^3 + as_1s_2 + bs_3.$$

Si $X_1 = X_2 = 1$, $X_i = 0$ pour $i > 2$, alors $p_3 = 2$, $s_1 = 2$, $s_2 = 1$, $s_3 = 0$. Si $X_1 = X_2 = X_3 = 1$, $X_i = 0$ pour $i > 3$, alors $p_3 = 3$, $s_1 = 3$, $s_2 = 3$, $s_3 = 1$. Du système obtenu

$$2 = 2^3 + a\cdot 2\cdot 1 + b\cdot 0,$$
$$3 = 3^3 + a\cdot 3\cdot 3 + b\cdot 1$$

on trouve $a = -3$, $b = 3$, c'est-à-dire $p_3 = s_1^3 - 3s_1s_2 + 3s_3$.

Les sommes de puissances $p_k(X_1, \dots, X_n)$ peuvent aussi être représentées sous forme de polynômes en $s_1, s_2, \dots, s_n$ à l'aide des

formules plus commodes, appelées *formules de Newton*:

$$p_k - p_{k-1}s_1 + p_{k-2}s_2 + \dots$$
$$\dots + (-1)^{k-1}p_1 s_{k-1} + (-1)^k k s_k = 0 \quad (9)$$

pour $1 \leqslant k \leqslant n$;

$$p_k - p_{k-1}s_1 + p_{k-2}s_2 + \dots$$
$$\dots + (-1)^{n-1}p_{k-n+1}s_{n-1} + (-1)^n p_{k-n}s_n = 0 \quad (10)$$

pour $k > n$.

Pour démontrer leur validité, servons-nous des relations évidentes

$$X_i^n - s_1 X_i^{n-1} + \dots + (-1)^{n-1} s_{n-1} X_i + (-1)^n s_n = 0,$$

obtenues par substitution $Y = X_i$ dans (2). En multipliant chacune de ces relations par X_i^{k-n} $(k \geqslant n)$:

$$X_i^k - s_1 X_i^{k-1} + \dots + (-1)^{n-1} s_{n-1} X_i^{k-n+1} + (-1)^n s_n X_i^{k-n} = 0$$

et effectuant ensuite la sommation sur i de 1 à n, nous obtenons non seulement la formule (10), mais aussi la formule (9) pour $k = n$ $(p_0 = X_1^0 + \dots + X_n^0 = n)$. Considérons ensuite le polynôme homogène symétrique $f_{k,n}$ de degré $k \leqslant n$ (ou de $-\infty$, si $f_{k,n} = 0$):

$$f_{k,n}(X_1, \dots, X_n) =$$
$$= p_k - p_{k-1}s_1 + \dots + (-1)^{k-1}p_1 s_{k-1} + \dots + (-1)^k k s_k.$$

Démontrons par récurrence sur $r = n - k$, que $f_{k,n}$ est identiquement nul. Pour $r = 0$, ce fait vient d'être établi. En posant $X_n = 0$ et en remarquant que les polynômes symétriques $(s_i)_0$, $(p_i)_0$ coïncident avec les polynômes s_i et p_i définis pour $n - 1$ indéterminées $X_1, \dots, X_{n-1}$ (voir (3) et (6)), nous obtenons l'égalité

$$f_{k,n}(X_1, \dots, X_{n-1}, 0) =$$
$$= (p_k)_0 - (p_{k-1})_0 (s_1)_0 + \dots + (-1)^{k-1} (p_1)_0 (s_{k-1})_0 +$$
$$+ (-1)^k k (s_k)_0 = f_{k, n-1}(X_1, \dots, X_{n-1}) = 0,$$

car $n - 1 - k = r - 1 < r$ et l'hypothèse de récurrence est applicable.

La relation $f_{k,n}(X_1, \dots, X_{n-1}, 0) = 0$ montre que le polynôme $f_{k,n}$ est divisible par X_n: $f_{k,n} = X_n f_1$. En utilisant le fait que $f_{k,n}$ est symétrique (voir la démonstration de la première partie du théorème 1), on obtient

$$f_{k,n}(X_1, \dots, X_n) = s_n(X_1, \dots, X_n) \cdot g(X_1, \dots, X_n),$$

ce qui, pourtant, ne peut avoir lieu que pour $g = 0$, car $\deg s_n = n$ et $\deg f_{k,n} = k < n$. Ainsi, $f_{k,n} = 0$, ce qui achève la démonstration de la formule (9).

4. Discriminant d'un polynôme.— Considérons dans l'anneau $P[X_1, \ldots, X_n]$ le polynôme

$$\Delta_n = \prod_{1 \leqslant j < i \leqslant n} (X_i - X_j),$$

qui peut évidemment être représenté sous la forme d'un déterminant de Vandermonde

$$\Delta_n = \begin{vmatrix} 1 & 1 & \ldots & 1 \\ X_1 & X_2 & \ldots & X_n \\ \cdot & \cdot & \cdot & \cdot \\ X_1^{n-1} & X_2^{n-1} & \ldots & X_n^{n-1} \end{vmatrix}. \qquad (11)$$

Puisque le déterminant est une fonction symétrique gauche de ses colonnes, $\pi(\Delta_n) = \varepsilon_\pi \Delta_n$, où ε_π est la signature de la permutation $\pi \in S_n$. Mais, dans ce cas, Δ_n^2 est un polynôme symétrique et, en vertu du théorème fondamental, il peut être représenté sous la forme d'un polynôme à fonctions symétriques élémentaires:

$$\Delta_n^2 = \Pi (X_i - X_j)^2 = \mathrm{Dis}(s_1, \ldots, s_n).$$

Le polynôme Dis en $s_1(X_1, \ldots, X_n), \ldots, s_n(X_1, \ldots, X_n)$ s'appelle *discriminant de la famille* $X_1, \ldots, X_n$. Ses coefficients sont évidemment contenus dans $\mathbb{Z}$. En substituant $x_i \in F$ à X_i, $i = 1, 2, \ldots, n$ (F est une extension quelconque du corps commutatif P), on peut parler du discriminant de la famille de n'importe quels n éléments du corps F. Si tous les $x_1, \ldots, x_n \in F$ ne sont pas distincts, le discriminant de cette famille s'annule, parce qu'au moins un des facteurs $x_i - x_j$ sera nul. C'est à sa propriété de discriminer ce cas, que le polynôme Dis doit son appellation de discriminant.

Un procédé bien commode permettant d'obtenir le discriminant est basé sur l'interprétation de Δ_n^2 en tant que produit du déterminant (11) par le déterminant transposé: $\Delta_n^2 = \Delta_n {}^t\Delta_n$ (rappelons que $\det {}^tA = \det A$ pour toute matrice carrée A).

En appliquant la règle de multiplication des matrices, on obtient tout de suite

$$\mathrm{Dis}(s_1, \ldots, s_n) = \begin{vmatrix} n & p_1 & p_2 & \ldots & p_{n-1} \\ p_1 & p_2 & p_3 & \ldots & p_n \\ p_2 & p_3 & p_4 & \ldots & p_{n+1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ p_{n-1} & p_n & p_{n+1} & \ldots & p_{2n-2} \end{vmatrix}, \qquad (12)$$

où p_k sont des sommes de puissances connues (6). En calculant p_k au moyen des formules récurrentes (9) et (10), on arrive à l'expression explicite de Dis $(s_1, \ldots, s_n)$. En particulier, $p_1 = s_1$,

$p_2 = s_1^2 - 2s_2$, de sorte que

$$\mathrm{Dis}\,(s_1, s_2) = \begin{vmatrix} 2 & s_1 \\ s_1 & s_1^2 - 2s_2 \end{vmatrix} = s_1^2 - 4s_2. \tag{13}$$

Soit maintenant donné un polynôme unitaire

$$f(X) = X^n + a_1X^{n-1} + \ldots + a_{n-1}X + a_n \in P[X],$$

possédant n zéros $c_1, \ldots, c_n$ dans P ou dans une extension quelconque F de ce corps. Comme on le sait des formules de Viète, $a_k = (-1)^k s_k(c_1, \ldots, c_n)$.

DÉFINITION. — *Le discriminant d'une famille de zéros* $c_1, \ldots, c_n$ *d'un polynôme* f *ou, ce qui est équivalent, la valeur du discriminant* Dis $(s_1, \ldots, s_n)$ *obtenue en substituant* $(-1)^k a_k$ à s_k, *s'appelle discriminant du polynôme* f *et se note* D (f). *Il s'appelle aussi discriminant de l'équation algébrique*

$$f(x) = x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n = 0. \tag{14}$$

Il est clair que D $(f) \in P$ (rappelons à ce sujet le corollaire du théorème 1). De la définition du discriminant on déduit aussi la proposition suivante:

PROPOSITION. — D $(f) = 0$ *si, et seulement si, l'équation* (14) *possède des racines multiples (au moins une racine de multiplicité* $k > 1$). ■

Compte tenu du corollaire 2 au théorème 5 du § 1, nous avons maintenant deux procédés qui n'exigent pas de considérer les extensions du corps de base P et permettent de déterminer si un polynôme $f \in P[X]$ possède des zéros multiples. Or, l'importance du discriminant ne se limite pas seulement à cela. Par exemple, la formule (13) appliquée au trinôme du second degré $f(X) = X^2 + aX + b$ à coefficients réels a, b, donne D $(f) = a^2 - 4b$, expression connue en Algèbre élémentaire. En particulier, c'est suivant le signe de D (f) que les racines de l'équation $x^2 + ax + b = 0$ sont réelles ou imaginaires conjuguées.

Calculons encore, à titre d'exemple, le discriminant de l'équation cubique dite incomplète

$$f(x) = x^3 + ax + b = 0. \tag{15}$$

Dans ce cas, $s_1 = 0$ et le calcul de p_k à l'aide des formules récurrentes donne $p_1 = s_1 = 0$, $p_2 = s_1^2 - 2s_2 = -2a$, $p_3 = s_1^3 - 3s_1s_2 + 3s_3 = -3b$, $p_4 = s_1^4 - 4s_1^2s_2 + 4s_1s_3 + 2s_2^2 = 2a^2$. Donc, d'après la formule (12), il vient

$$\mathrm{D}\,(f) = \begin{vmatrix} 3 & 0 & -2a \\ 0 & -2a & -3b \\ -2a & -3b & 2a^2 \end{vmatrix} = -4a^3 - 27b^2. \tag{16}$$

L'expression pour D (f) prend une forme plus compliquée (par rapport à (16)) dans le cas de l'équation cubique complète $x^3 + a_1x^2 + a_2x + a_3 = 0$, pourtant il est possible de ne pas le considérer comme le montre le raisonnement général suivant.

Passons de l'argument x à $y = x + \frac{a_1}{n}$. En portant $x = y - \frac{a_1}{n}$ dans l'équation (14) et en se servant de la formule du binôme, on trouve que

$$g(y) = f\left(y - \frac{a_1}{n}\right) = y^n + ay^{n-2} + \ldots = 0, \tag{17}$$

c'est-à-dire que dans la nouvelle équation le coefficient de y^{n-1} est nul. Connaissant la racine y_0 de l'équation (17), on calcule sans peine la racine $x_0 = y_0 - \frac{a_1}{n}$ de l'équation initiale (14). Aussi, sans restreindre la généralité, peut-on poser $a_1 = 0$.

Si l'on essaie de trouver une formule générale, donnant la solution de l'équation (15) (ce qu'ont réussi à faire les mathématiciens de Moyen Age Spicione dal Ferro, Cardan et autres), on devra inévitablement mettre en jeu le discriminant (16) (voir formules (2) du chap. 1, § 2).

5. Résultant.— La propriété essentielle de D (f), énoncée dans la proposition du numéro précédent, peut aussi être interprétée comme un indice de ce que le polynôme f et sa dérivée f' possèdent des zéros communs (ou des facteurs communs). C'est l'algorithme d'Euclide qui est, en fin de compte, à la base de cet indice. Cela nous permet de supposer qu'il existe un critère analogue, permettant de déterminer directement d'après les coefficients de deux polynômes quelconques f, $g \in P[X]$ si ces polynômes possèdent ou non un facteur commun.

Ainsi, soient

$$f(X) = a_0X^n + a_1X^{n-1} + \ldots + a_{n-1}X + a_n,$$
$$g(X) = b_0X^m + b_1X^{m-1} + \ldots + b_{m-1}X + b_m$$

deux polynômes à coefficients dans un corps commutatif P. Ici, $n > 0$, $m > 0$, mais il n'est pas exclu que $a_0 = 0$ ou $b_0 = 0$.

DÉFINITION. — *On appelle résultant* Res (f, g) *des polynômes f et g un polynôme homogène* (*une fonction polynomiale homogène*) *à coefficients de f et g* (*de degré m par rapport à $a_0, \ldots, a_n$ et de degré n par rapport à $b_0, \ldots, b_m$*) *de la forme*

$$\operatorname{Res}(f, g) = \left|\begin{matrix} a_0 & a_1 & \ldots & a_n & & & \\ & a_0 & a_1 & \ldots & a_n & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & a_0 & a_1 & \ldots & a_n \\ b_0 & b_1 & \ldots & b_m & & & \\ & b_0 & b_1 & \ldots & b_m & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & b_0 & b_1 & \ldots & b_m \end{matrix}\right| \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} m \text{ lignes} \\ \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} n \text{ lignes} \end{matrix}$$

Cette définition contient une certaine affirmation concernant les degrés du résultant en tant que polynôme. Or, cette affirmation découle directement des propriétés des déterminants: si, dans les m premières lignes, on remplace a_i par ta_i, on obtient Res $(tf, g) = t^m$ Res (f, g), après quoi il ne reste qu'à se reporter à l'exercice 3 du chap. 5, § 2.

Etablissons maintenant les propriétés essentielles du résultant.

Res 1. *Res $(f, g) = 0$ si, et seulement si, $a_0 = 0 = b_0$ ou bien f et g possèdent dans $P[X]$ un facteur commun de degré > 0.*

Assurons-nous d'abord que la condition « $a_0 = 0 = b_0$ ou bien les polynômes f et g possèdent dans $P[X]$ un facteur commun de degré > 0 » est satisfaite si, et seulement si, il existe des polynômes f_1, g_1 simultanément non nuls, tels que l'on ait

$$fg_1 + f_1g = 0, \ \deg f_1 < n, \ \deg g_1 < m. \qquad (18)$$

En effet, soit $h =$ P.G.C.D. (f, g), $\deg h > 0$. Alors, $f = hf_1$; $g = -hg_1$ et donc $fg_1 + gf_1 = 0$. De plus, $\deg f_1 < n$, $\deg g_1 < m$, si bien que la condition (18) est satisfaite. Pour $a_0 = 0 = b_0$, nous pouvons poser $f_1 = f$, $g_1 = -g$.

Réciproquement, en supposant que P.G.C.D. $(f, g) = 1$, la condition (18) étant satisfaite, nous obtenons du fait que $P[X]$ est factoriel (voir chap. 5, § 3), l'implication $fg_1 = -gf_1 \Rightarrow f \mid f_1$, $g \mid g_1$. Par conséquent, $\deg f < n$, $\deg g < m$, d'où $a_0 = 0 = b_0$.

Démontrons maintenant l'équivalence des conditions (18) et Res $(f, g) = 0$. En posant

$$f_1 = c_0X^{n-1} + c_1X^{n-2} + \ldots + c_{n-1},$$
$$g_1 = d_0X^{m-1} + d_1X^{m-2} + \ldots + d_{m-1}$$

et calculant, d'après les règles formelles, les coefficients du polynôme $fg_1 + f_1g$ de degré $\leqslant n + m - 1$, nous écrirons la condition (18) sous la forme d'un système homogène carré d'équations linéaires à $(n + m)$ inconnues $d_0, d_1, \ldots, d_{m-1}, c_0, c_1, \ldots, c_{n-1}$:

$$\begin{array}{ll} a_0d_0 + \ldots\ldots\ldots\ldots\ldots\ldots + b_0c_0 \ldots\ldots\ldots = 0, & \\ a_1d_0 + a_0d_1 + \ldots\ldots\ldots\ldots + b_1c_0 + b_0c_1 \ \ldots = 0, & \\ a_2d_0 + a_1d_1 + a_0d_2 \ldots \quad + b_2c_0 + b_1c_1 + b_0c_2 = 0, & (19) \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \end{array}$$

Le déterminant de la matrice du système (19) (plus exactement, le déterminant de la matrice transposée) coïncide justement avec Res (f, g). Il s'ensuit que le système (19) admet une solution non nulle si, et seulement si, Res $(f, g) = 0$. Or, toute solution non nulle entraîne l'existence d'un couple de polynômes f_1, g_1 qui satisfont à la condition (18). ∎

Res 2. *Soient f et g deux polynômes qui se scindent dans $P[X]$ en facteurs linéaires*

$$f(X) = a_0 (X - \alpha_1) \ldots (X - \alpha_n),$$
$$g(X) = b_0 (X - \beta_1) \ldots (X - \beta_m).$$

Alors, on a

$$\operatorname{Res}(f, g) = a_0^m \prod_{i=1}^{n} g(\alpha_i) = (-1)^{mn} b_0^n \prod_{j=1}^{m} f(\beta_j) = a_0^m b_0^n \prod_{i,j} (\alpha_i - \beta_j).$$

Démonstration. — Il est clair que, si elles sont vraies, les formules indiquées ci-dessus doivent avoir un caractère universel, indépendant des types particuliers des polynômes f, g. Cette « philosophie » assez simple dont la nature ne sera pas examinée ici, permet de nous borner à étudier le « cas général » où, disons, tous les $g(\alpha_1), \ldots, g(\alpha_n)$ et tous les $f(\beta_1), \ldots, f(\beta_m)$ sont deux à deux distincts.

Puisque $\operatorname{Res}(g, f) = (-1)^{mn} \operatorname{Res}(f, g)$ (voir définition), il suffit de vérifier la relation $\operatorname{Res}(f, g) = a_0^m \prod g(\alpha_i)$. A cet effet, introduisons une nouvelle indéterminée Y et considérons les polynômes $f(X)$, $g(X) - Y$ sur le corps $P(Y)$ des fractions rationnelles. De la définition du résultant, où il convient de remplacer b_m par $b_m - Y$, on déduit que

$$\operatorname{Res}(f, g - Y) = (-1)^n a_0^m Y^n + \ldots + \operatorname{Res}(f, g)$$

est un polynôme de degré n par rapport à Y, dont le coefficient dominant est $(-1)^n a_0^m$ et le terme constant $\operatorname{Res}(f, g)$. Les polynômes $f(X)$ et $g(X) - g(\alpha_i)$ ayant le zéro commun α_i sont divisibles par $X - \alpha_i$. En raison de la propriété Res 1, on a $\operatorname{Res}(f, g - g(\alpha_i)) = 0$.

D'après le théorème de Bézout, le polynôme $\operatorname{Res}(f, g - Y)$ doit être divisible par $g(\alpha_i) - Y$, $1 \leqslant i \leqslant n$. Puisque tous les $g(\alpha_i)$ sont distincts, on a $\operatorname{Res}(f, g - Y) = a_0^m \prod_{i=1}^{n} (g(\alpha_i) - Y)$. Pour $Y = 0$, on obtient l'expression requise. ■

Etendons la définition du discriminant, donnée au point 4, au cas des polynômes non unitaires, en posant

$$D(f) = a_0^{2n-2} \prod_{1 \leqslant j < i \leqslant n} (\alpha_i - \alpha_j)^2 = \left[a_0^{n-1} \prod_{j<i} (\alpha_i - \alpha_j)\right]^2, \quad a_0 \neq 0.$$

Res 3. *On a la formule suivante*

$$D(f) = (-1)^{\frac{n(n-1)}{2}} a_0^{-1} \operatorname{Res}(f, f'). \tag{20}$$

En effet, suivant Res 2, on a

$$\operatorname{Res}(f, f') = a_0^{n-1} \prod_{i=1}^{n} f'(\alpha_i).$$

Or,

$$f'(\alpha_i) = a_0 \prod_{j \neq i} (\alpha_i - \alpha_j),$$

ce qui est une simple conséquence de la substitution $X = \alpha_i$ dans l'expression générale

$$f'(X) = a_0 \sum_{i=1}^{n} \prod_{j \neq i} (X - \alpha_j),$$

obtenue par dérivation du produit $f(X) = a_0 \prod_{j=1}^{n} (X - \alpha_j)$. Ainsi,

$$\text{Res}(f, f') = a_0^{2n-1} \prod_{i=1}^{n} \prod_{j \neq i} (\alpha_i - \alpha_j) =$$

$$= a_0 (-1)^{\frac{n(n-1)}{2}} a_0^{2n-2} \prod_{j<i} (\alpha_i - \alpha_j)^2 = a_0 (-1)^{\frac{n(n-1)}{2}} \text{D}(f). \quad \blacksquare$$

La formule (20) fournit une expression explicite du déterminant.

EXERCICES

1. Etant donné un nombre premier p, montrer, en se servant des formules de Newton (9), (10), que

$$\sum_{i=1}^{p-1} i^m \equiv \begin{cases} -1 \pmod p, & \text{si } m \text{ est divisible par } p-1, \\ 0 \pmod p, & \text{si } m \text{ n'est pas divisible par } p-1. \end{cases}$$

2. Soient c_1, c_2, c_3 les racines complexes du polynôme $X^3 - X + 1$. Que peut-on dire de l'extension $\mathbb{Q}(c_1^{99} + c_2^{99} + c_3^{99})$?

3. On dit qu'un polynôme $f(X_1, \ldots, X_n)$ sur un corps commutatif P de caractéristique $\neq 2$ est *antisymétrique* (ou *alterné*) si $(\pi f)(X_1, \ldots, X_n) = \varepsilon_\pi f(X_1, \ldots, X_n)$, $\forall \pi \in S_n$ (comme toujours, ε_π est la signature de la permutation). Comme exemple de polynôme antisymétrique on peut indiquer $\Delta_n = \prod_{j<i} (X_i - X_j)$. Montrer que tout polynôme antisymétrique $f \in P[X_1, \ldots, X_n]$ est de la forme $f = \Delta_n \cdot g$, où g est un polynôme symétrique. (Indication. Considérer f comme polynôme en X_n à coefficients dans $P[X_1, \ldots, X_{n-1}]$. Porter attention sur le fait qu'étant antisymétrique, $f = 0$ pour $X_n = X_{n-1}$ et donc f est divisible par $X_n - X_{n-1}$.)

4. En utilisant la propriété Res 2 et le fait qu'il existe un corps de décomposition du polynôme (voir théorème 2 dans le paragraphe suivant), montrer que

$$\text{Res}(fg, h) = \text{Res}(f, h) \cdot \text{Res}(g, h).$$

5. En utilisant les résultats de l'exercice 4 et la propriété Res 3, obtenir la formule

$$\text{D}(fg) = \text{D}(f)\, \text{D}(g)\, [\text{Res}(f, g)]^2.$$

6. Quelle est la valeur du résultant $\text{Res}(f(X), X - a)$?

7. Montrer que $\text{D}(X^n + a) = (-1)^{\frac{n(n-1)}{2}} n^n a^{n-1}$.

8. Soit $f(X) = X^{n-1} + X^{n-2} + \ldots + 1$. A partir de la relation $X^n - 1 = (X - 1) f(X)$ et des exercices précédents, montrer que $D(f) = (-1)^{\frac{(n-1)(n-2)}{2}} n^{n-2}$.

§ 3. Le corps $\mathbb{C}$ est algébriquement clos

1. Enoncé du théorème fondamental.— Soient P un corps commutatif et f un polynôme quelconque sur P. Comme il a été déjà dit au § 1, n° 2, le comportement de la fonction polynomiale $\tilde{f}$: $P \to P$ associée à f dépend fortement du corps P. En particulier, $\operatorname{Im} \tilde{f} = P$ si $\deg f > 0$ et on peut appliquer à P la définition suivante:

Définition. — *Un corps P est dit algébriquement clos (ou parfois algébriquement fermé) si tout polynôme de l'anneau $P[X]$ se décompose en facteurs linéaires.*

On peut donner un énoncé équivalent: *un corps P est algébriquement clos si les seuls polynômes irréductibles sur P sont les polynômes de degré 1 (des polynômes linéaires).*

Si tout polynôme $f \in P[X]$ admet au moins une racine dans P, le corps P est algébriquement clos.

En effet, on a alors $f(X) = (X - a) h(X)$, $a \in P$, $h \in P[X]$. Or, par hypothèse, le polynôme h à coefficients dans P admet, lui aussi, au moins une racine, c'est-à-dire $h(X) = (X - b) r(X)$, $b \in P$, $r \in P[X]$. En réitérant cette opération, nous obtiendrons finalement la décomposition complète de f en facteurs linéaires. Puisque f était un polynôme arbitraire, le corps P satisfait à la définition d'un corps algébriquement clos.

Bien que l'assertion que tout corps P admet une extension $\tilde{P} \supset P$ qui est un corps algébriquement clos (*théorème de Steinitz*), soit vraie, il est tout de même difficile, en tout cas les premiers temps, d'assimiler non seulement la construction d'une extension algébriquement close, mais l'idée même d'une telle extension. Il est d'autant plus agréable de constater que nous avons un exemple magnifique et très important de corps algébriquement clos, comme l'énonce le théorème que l'on appelle « théorème fondamental de l'algèbre »:

Théorème 1.— *Le corps $\mathbb{C}$ des nombres complexes est algébriquement clos.*

Enonçons encore une fois cette assertion fondamentale, mais maintenant en termes de racines:

Un polynôme arbitraire $f(X)$ de degré $n \geqslant 1$ à coefficients complexes (ou réels) possède exactement n racines complexes, comptées avec leur multiplicité.

Le titre de « fondamental » a été conféré au théorème 1 encore à l'époque où la résolution des équations algébriques était une des occupations principales des algébristes. De nos jours, le théorème 1 se range dans la catégorie des assertions ordinaires, bien qu'importantes.

La première démonstration rigoureuse du « théorème fondamental » a été donnée par Gauss en 1799. Beaucoup de variantes de démonstration, qui se distinguent entre elles par le degré d'algébricité, si l'on peut dire ainsi, ont été proposées par la suite. La nécessité de s'appuyer sur la propriété de continuité des corps $\mathbb{R}$ et $\mathbb{C}$ (autrement dit sur leur topologie) se manifeste sous une forme ou sous une autre ; il existe même une démonstration tout à fait non algébrique et bien courte qui se base sur une notion assez profonde de fonction analytique d'une variable complexe. La démonstration que nous donnons ci-dessous est, par son esprit, la plus algébrique de toutes qui nous soient accessibles. La plus naturelle serait peut-être la démonstration qui utilise les moyens fournis par la théorie de Galois, mais nous nous contenterons de cette seule mention.

La partie non algébrique de la démonstration du théorème 1 est représentée par deux lemmes suivants:

LEMME 1 (lemme sur le module du terme de plus haut degré). — *Soit*

$$f(X) = a_0X^n + a_1X^{n-1} + \ldots + a_{n-1}X + a_n \qquad (1)$$

un polynôme de degré $n \geqslant 1$ *à coefficients complexes arbitraires. Alors, pour l'application polynomiale* $z \mapsto f(z)$ *du corps* $\mathbb{C}$ *dans lui-même, on peut indiquer un nombre positif* $r \in \mathbb{R}$ *tel que pour* $|z| > r$ *on ait l'inégalité*

$$|a_0z^n| > |a_1z^{n-1} + \ldots + a_{n-1}z + a_n|.$$

DÉMONSTRATION. — Posons $A = \max(|a_1|, \ldots, |a_n|)$ et $r = \frac{A}{|a_0|} + 1$. Si l'on prend $|z| > r \geqslant 1$, on obtient $|a_0| > \frac{A}{|z|-1} - 1$, d'où, suivant les règles relatives aux opérations sur les modules des nombres complexes (voir chap. 5, § 1), on a:

$$|a_0z^n| = |a_0|\,|z|^n > \frac{A|z|^n}{|z|-1} > \frac{A(|z|^n-1)}{|z|-1} =$$
$$= A(|z|^{n-1} + \ldots + |z| + 1) \geqslant |a_1|\,|z|^{n-1} + \ldots$$
$$+ |a_{n-1}|\,|z| + |a_n| = |a_1z^{n-1}| + \ldots + |a_{n-1}z| + |a_n| \geqslant$$
$$\geqslant |a_1z^{n-1} + \ldots + a_{n-1}z + a_n|. \;\blacksquare$$

COROLLAIRE. — *Soit* (1) *un polynôme de degré* $n \geqslant 1$ *à coefficients réels. Alors, pour tout* $x \in \mathbb{R}$, *suffisamment grand en valeur absolue, le signe de* $f(x)$ (*du nombre réel*) *coïncide avec le signe du « terme de plus haut degré »* a_0x^n. ■

LEMME 2. — *Un polynôme de degré impair à coefficients réels admet au moins une racine réelle.*

DÉMONSTRATION. — Puisque n est impair, le terme de plus haut degré a_0x^n de l'application polynomiale $\tilde{f}$: $\mathbb{R} \to \mathbb{R}$ prendra des signes différents pour des $x \in \mathbb{R}$ positifs et négatifs. En prenant ces valeurs de x suffisamment grandes en valeur absolue, nous pouvons affirmer, en vertu du corollaire au lemme 1, que les $f(x)$ seront aussi de signes différents. Si, par exemple $a_0 > 0$, on aura $f(-r) < 0$ et $f(r) > 0$, où r est un nombre réel pris de la démonstration du lemme 1. Comme on l'apprend en Analyse mathématique (d'ailleurs, ce qui est facile à montrer directement), l'application polynomiale $\tilde{f}$ est continue (autrement dit, une fonction rationnelle entière $x \mapsto f(x)$ est continue). La fonction continue f possède la propriété de prendre sur l'intervalle $-r \leqslant x \leqslant r$ toute valeur intermédiaire entre $f(-r)$ et $f(r)$. En particulier, pour un c quelconque tel que $|c| \leqslant r$, on aura $f(c) = 0$. Le même raisonnement est applicable à $a_0 < 0$. ∎

Nous terminons nos raisonnements non algébriques par cette assertion qui est géométriquement et intuitivement claire. Nous développerons l'étape suivante de la démonstration dans un contexte qui n'a pas de rapport direct avec le corps $\mathbb{C}$, et considérerons une structure présentant un intérêt indépendant.

2. Corps de décomposition d'un polynôme. — Comme il arrive bien souvent, un nouveau coup d'œil jeté sur un exemple bien connu permet de mieux comprendre cet exemple et de passer à des généralisations raisonnables. Rappelons la réalisation du corps $\mathbb{C}$ sous la forme de l'anneau quotient $\mathbb{R}[X]/(x^2+1)\,\mathbb{R}[X]$ (chap. 5, § 2, théorème 6). En y remplaçant $\mathbb{R}$ par un corps commutatif quelconque P, et X^2+1 par un polynôme arbitraire $f \in P[X]$, nous obtenons un « anneau des classes résiduelles modulo, (f) » ou, ce qui revient au même, un anneau quotient $P[X]/(f)$, où $(f) = f \cdot P[X]$ est un idéal de $P[X]$. L'idéal (f) se compose de tous les polynômes divisibles par f et, en vertu du corollaire au théorème 5 du chap. 5, § 2, est l'idéal le plus général de $P[X]$. L'analogie qui existe entre les anneaux $\mathbb{Z}$ et $P[X]$ s'étend aux anneaux correspondants des classes résiduelles $Z_n = \mathbb{Z}/(n)$ et $P[X]/(f)$. Il est utile de reprendre les étapes principales de la construction de Z_n décrites au chapitre 4, § 4.

Les éléments de l'anneau quotient $P[X]/(f)$ sont des classes résiduelles $\bar{g} = g + (f)$ dont chacune peut être représentée sous la forme $r + (f)$, où $\deg r < \deg f$. Tout comme dans le cas de $\mathbb{Z}$, la démonstration est fournie, cette fois aussi, par la division euclidienne: si $g = qf + r$, on a $g + (f) = r + qf + (f) = r + (f)$, puisque

$qf \in (f)$. Il est facile de voir que les éléments $\bar{a}$, $a \in P$, forment dans $P[X]/(f)$ un sous-anneau isomorphe au corps P. La réductibilité du polynôme f sur P, c'est-à-dire la possibilité de sa représentation sous la forme $f = f_1 f_2$, où $f_i \in P[X]$ et $0 < \deg f_i < \deg f$, entraîne l'existence dans $P[X]/(f)$ de diviseurs non triviaux de zéro; à savoir: $\bar{f}_i \neq \bar{0}$, $i = 1, 2$, mais $\bar{f}_1 \bar{f}_2 = \overline{f_1 f_2} = \bar{f} = \bar{0}$.

Supposons maintenant que f soit un polynôme irréductible. Si $\deg r < \deg f$ $(r \neq 0)$, alors P.G.C.D. $(r, f) = 1$ et $ur + vf = 1$ pour certains $u, v \in P[X]$ (voir chap. 5, § 3, théorème 3). En d'autres termes

$$\{r + (f)\}\{u + (f)\} = ru + (f) = 1 - vf + (f) = 1 + (f),$$

d'où

$$\bar{r}\bar{u} = \overline{ru} = \bar{1}.$$

Cela signifie que tout élément $\bar{r} \neq \bar{0}$ admet dans $P[X]/(f)$ un inverse $\bar{u} = \bar{r}^{-1}$. Cette remarque montre que dans le cas où le polynôme f est irréductible, l'anneau quotient $P[X]/(f)$ est un corps contenant un sous-corps isomorphe à P.

Portons notre attention sur l'élément spécial $\bar{X} \in P[X]/(f)$. Quels que soient $a_0, a_1, \ldots, a_n \in P$, on a

$$\sum_{k=0}^{m} \bar{a}_k \bar{X}^k = \sum_k \{a_k + (f)\}\{X + (f)\}^k =$$
$$= \sum_k \{a_k + (f)\}\{X^k + (f)\} = \{\sum_k a_k X^k\} + (f) = \overline{\sum_k a_k X^k}.$$

En bref, si $g(Y) = \sum a_k Y^k \in P[Y]$, alors $g(\bar{X}) = \overline{g(X)}$. L'écriture $g(\bar{X})$ a, certes, son sens lorsqu'on identifie P et le corps qui est contenu dans $P[X]/(f)$ et est isomorphe à P. En particulier

$$f(\bar{X}) = \overline{f(X)} = f + (f) = (f) = \bar{0},$$

c'est-à-dire l'élément $\bar{X} \in P[X]/(f)$ est un zéro du polynôme (f).

Ainsi, on peut énoncer le théorème suivant:

THÉORÈME 2.— *L'anneau des classes résiduelles (l'anneau quotient) $P[X]/(f)$ est un corps si, et seulement si, f est un polynôme irréductible sur P.* ■

COROLLAIRE.— *Pour tout polynôme $f(X)$ irréductible sur un corps P, il existe une extension $F \supset P$ dans laquelle $f(X)$ admet au moins un zéro. Pour F, on peut prendre le corps isomorphe à $P[X]/(f)$.* ■

Conformément à la terminologie en vigueur, on convient de dire que l'extension F est obtenue par adjonction à P d'un zéro c du polynôme f: $F = P(c)$. Ceci étant, $f(X) = (X - c)\, g(X)$,

où $g \in F[X]$. Nous avons donc une possibilité réelle de construire une extension du corps P dans laquelle le polynôme f se décompose complètement en facteurs linéaires.

DÉFINITION.— *Soient P un corps et f un polynôme unitaire (non nécessairement irréductible) de degré n de $P[X]$. Alors, on dit que l'extension $F \supset P$ est un corps de décomposition de f sur P si $f(X) = (X - c_1) \ldots (X - c_n)$ dans $F[X]$ et $F = P(c_1, \ldots, c_n)$, c'est-à-dire F est obtenue à partir de P par adjonction des racines $c_1, \ldots, c_n$ du polynôme f.*

THÉORÈME 3. — *Pour tout polynôme unitaire $f \in P[X]$ de degré $n > 0$, il existe au moins un corps de décomposition.*

DÉMONSTRATION. — La condition pour que le polynôme soit unitaire n'est pas trop nécessaire, elle n'est utilisée que pour la commodité. Soit

$$f(X) = f_1(X) \ldots f_r(X)$$

la décomposition de f en facteurs unitaires irréductibles dans $P[X]$. En vertu du corollaire au théorème 2 il existe une extension $P_1 \supset P$ qui contient au moins une racine du polynôme f_1. Cette racine c_1 sera évidemment aussi racine de f. Supposons qu'on ait déjà trouvé une extension $P_k \supset \ldots \supset P_1 \supset P$ sur laquelle f possède la décomposition

$$f(X) = (X - c_1) \ldots (X - c_k)\, g_1(X) \ldots g_s(X)$$

avec k facteurs linéaires (non nécessairement distincts), $k < n$. En appliquant de nouveau le corollaire du théorème 2 au corps P_k et au polynôme unitaire irréductible $g_1 \in P_k[X]$, nous construirons un corps $P_{k+1} \supset P_k$ sur lequel il sera possible de scinder le polynôme $g_1(X)$ et, par conséquent le polynôme $f(X)$ en obtenant le facteur linéaire $X - c_{k+1}$, avec $c_{k+1} \in P_{k+1}$. Par application répétée de ce procédé nous obtiendrons une décomposition complète de f en facteurs linéaires sur une certaine extension $P_n \supset P$. Soit P_n, soit son sous-corps quelconque F sera justement le corps de décomposition de f. Il n'est pas exclu que F peut coïncider avec P. ■

La démonstration du théorème 3 comporte trop d'arbitraire pour qu'on puisse parler de l'unicité du corps de décomposition du polynôme f; et bien qu'en réalité le corps de décomposition soit déterminé de façon unique, à un isomorphisme près, il est assez difficile de le démontrer. Pour le moment, nous n'aurons pas besoin de cette propriété supplémentaire des corps de décomposition.

EXEMPLES. 1) Le corps quadratique $\mathbb{Q}(\sqrt{d})$ est un corps de décomposition du polynôme $X^2 - d$.

2) Si l'on adjoint à Z_2 la racine θ du polynôme irréductible $X^2 + X + 1$, on obtient le corps $Z_2(\theta) = \{0, 1, \theta, 1 + \theta\}$ à quatre éléments, qui est isomorphe tant au corps $Z_2[X]/(X^2 + X + 1)$ qu'au corps GF (4) (voir chap. 4, § 4,

n° 6). Remarquons que $X^2 + X + 1 = (X - \theta)(X - \theta^2)$, c'est-à-dire que $Z_2(\theta)$ est un corps de décomposition du polynôme $X^2 + X + 1$.

3) Le polynôme $X^2 + 1$ est irréductible non seulement sur $\mathbb{R}$, lorsque son corps de décomposition sera $\mathbb{C}$, mais aussi sur certains autres corps, par exemple sur Z_3. Soit $\theta^2 = -1$ (si l'on veut, $\theta = X + (X^2 + 1) Z_3[X]$ est un élément du corps des classes résiduelles $Z_3[X]/(X^2+1)$). Puisque $X^2 + 1 = (X - \theta) \times \times (X - 2\theta)$, le corps $Z_3(\theta) = \{a + b\theta \mid a, b \in Z_3\}$ est un corps de décomposition de $X^2 + 1$ sur Z_3. Signalons en passant que $Z_3(\theta)$ est isomorphe au corps des matrices $\begin{Vmatrix} a & b \\ -b & a \end{Vmatrix}$, $a, b \in Z_3$ (voir exercice 14 du chap. 4, § 4). L'application correspondante est: $a + b\theta \longmapsto a \begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix} + b \begin{Vmatrix} 0 & 1 \\ -1 & 0 \end{Vmatrix}$. Portons notre attention sur le fait que $Z_3(\theta)^* = \langle\lambda\rangle$, $\lambda = 1 + \theta$, $\lambda^2 = -\theta$, $\lambda^3 = 1 - \theta$, $\lambda^4 = -1$, $\lambda^5 = -1 - \theta$, $\lambda^6 = \theta$, $\lambda^7 = -1 + \theta$, $\lambda^8 = 1$, c'est-à-dire que le groupe multiplicatif du corps $Z_3(\theta)$ est non seulement abélien mais aussi cyclique.

4) D'après le critère d'Eisenstein, le polynôme $X^3 - 2$ est irréductible sur $\mathbb{Q}$. Toute racine de $X^3 - 2$ n'étant pas réelle, $\mathbb{Q}(\sqrt[3]{2})$ ne peut pas être un corps de décomposition. Le corps de décomposition de $X^3 - 2$ est $\mathbb{Q}(\sqrt[3]{2}, \varepsilon)$, où ε est la racine primitive cubique de l'unité:

$$X^3 - 2 = (X - \sqrt[3]{2})(X - \varepsilon\sqrt[3]{2})(X - \varepsilon^2\sqrt[3]{2}).$$

3. Démonstration du théorème fondamental. — Du numéro précédent, nous n'aurons besoin que de la définition du corps de décomposition et de l'assertion du théorème 3.

Conformément à la remarque faite immédiatement après la définition du corps algébriquement clos, il est nécessaire d'établir que le polynôme (1) admet au moins une racine complexe. Supposons d'abord que tous les coefficients de ce polynôme soient réels et admettons sans restreindre la généralité que $a_0 = 1$, $a_n \neq 0$. Soit

$$\deg f = 2^m n_0,$$

où n_0 est un entier impair. Dans le cas où $m = 0$, le polynôme f possède, d'après le lemme 2, une racine qui est même réelle. En raisonnant par récurrence sur m, considérons le théorème démontré pour tous les polynômes à coefficients réels, dont le degré est de la forme $2^{m'} n_0'$, avec $m' \leqslant m - 1$ (le facteur impair n_0' n'est soumis à aucune limitation).

Considérons le corps de décomposition F du polynôme $(X^2 + 1) f(X)$, qui existe en raison du théorème 3 et contient $\mathbb{C}$ comme sous-corps. Soient $u_1, u_2, \ldots, u_n$ les racines du polynôme f dans F. Considérons dans F les éléments

$$v_{ij} = u_i u_j + a(u_i + u_j), \quad 1 \leqslant i < j \leqslant n, \tag{2}$$

où a est un nombre réel fixe quelconque. Il faudrait écrire $v_{ij}(a)$, mais pour ne pas compliquer les notations nous ne le ferons pas. Le nombre n' d'éléments de la forme (2) est égal à

$$n' = \binom{n}{2} = \frac{n(n-1)}{2} = \frac{2^m n_0 (2^m n_0 - 1)}{2} = 2^{m-1} n_0', \tag{3}$$

où n_0' est un entier impair

Le polynôme

$$f_a(X) = \prod_{1 \leqslant i < j \leqslant n} (X - v_{ij}) = X^{n'} + b_1 X^{n'-1} + \ldots + b_{n'}$$

de l'anneau $F[X]$ est de degré n' et, par définition, tous les éléments (2) sont ses racines. D'après les formules de Viète (12) du § 1, les coefficients $b_1, \ldots, b_{n'}$ du polynôme $f_a(X)$ seront, au signe près, des fonctions symétriques élémentaires s_k de v_{ij}. En introduisant dans $s_k(v_{12}, v_{13}, \ldots, v_{n-1,n})$ les expressions des éléments v_{ij} par $u_1, \ldots, u_n$, on obtient la fonction

$$h_k(u_1, \ldots, u_n) = s_k(\ldots, u_i u_j + a(u_i + u_j), \ldots),$$
$$k = 1, \ldots, n',$$

qui est aussi symétrique. En effet, pour toute permutation $\pi \in S_n$ (S_n est le groupe symétrique de degré n) on a

$$\hat{\pi} v_{ij} = u_{\pi(i)} u_{\pi(j)} + a(u_{\pi(i)} + u_{\pi(j)}) = v_{\pi(i),\, \pi(j)}$$

(ou $v_{\pi(j),\, \pi(i)}$, si $\pi(i) > \pi(j)$), de sorte que π induit une permutation $\hat{\pi}$ sur l'ensemble des éléments de la forme (2). Etant symétrique, $s_k(v_{12}, v_{13}, \ldots, v_{n-1,n})$ ne change pas lors de la permutation des arguments et donc

$$(\pi h_k)(u_1, \ldots, u_n) = s_k(\hat{\pi} v_{12}, \hat{\pi} v_{13}, \ldots, \hat{\pi} v_{n-1,n}) =$$
$$= s_k(v_{12}, v_{13}, \ldots, v_{n-1,n}) = h_k(u_1, \ldots, u_n).$$

Remarquons que $h_k(u_1, \ldots, u_n)$ est, pour $X_i = u_i$, $i = 1, \ldots, n$, la valeur du polynôme symétrique $h_k(X_1, \ldots, X_n)$ à coefficients réels qui ne dépendent que de $a \in \mathbb{R}$.

En raison du théorème fondamental sur les polynômes symétriques (§ 2, théorème 1), il existe un polynôme $g_k(Y_1, \ldots, Y_n)$ à coefficients réels, tel que l'on ait $h_k(X_1, \ldots, X_n) = g_k(s_1(X_1, \ldots, X_n), \ldots, s_n(X_1, \ldots, X_n))$. Donc,

$$(-1)^k b_k = h_k(u_1, \ldots, u_n) =$$
$$= g_k(s_1(u_1, \ldots, u_n), \ldots, s_n(u_1, \ldots, u_n)) =$$
$$= g_k(-a_1, \ldots, (-1)^n a_n) \in \mathbb{R}$$

(rappelons que a_i sont les coefficients du polynôme unitaire $f \in \mathbb{R}[X]$ que nous considérons).

Ainsi, les coefficients b_k du polynôme $f_a(X)$ se sont avérés réels pour tout $a \in \mathbb{R}$. Puisque $\deg f_a = n' = 2^{m-1} n_0'$ (voir (3)), le polynôme f_a possède, par hypothèse de récurrence, au moins une racine complexe qui doit coïncider bien sûr avec l'un des éléments v_{ij}. En faisant varier le paramètre $a \in \mathbb{R}$, nous obtiendrons d'autres polynômes $f_a(X)$ à coefficients réels. Or, à chacun de ces polynômes il correspond un couple d'indices $i < j$ (dépendant de a), tel que

l'élément $v_{ij} = u_i u_j + a(u_i + u_j) \in F$ soit contenu dans le sous-corps $\mathbb{C}$ du corps F. Puisque les couples d'indices $i < j$ distincts ne sont qu'au nombre de $\binom{n}{2}$, alors que l'ensemble des nombres réels $a \in \mathbb{R}$ est infiniment grand, il existe deux nombres réels distincts a, a' auxquels il correspond un seul et même couple d'indices, disons $i = 1$, $j = 2$ (ce n'est que la question de numération des racines $u_1, \ldots, u_n$) pour lesquels

$$\begin{gathered} u_1u_2 + a(u_1 + u_2) = c, \\ u_1u_2 + a'(u_1 + u_2) = c', \; a \neq a', \end{gathered} \tag{4}$$

seront des nombres complexes. Du système d'équations (4) on déduit que

$$u_1 + u_2 = \frac{c - c'}{a - a'}, \quad u_1u_2 = c - a\,\frac{c - c'}{a - a'}$$

appartiennent, eux aussi, au corps $\mathbb{C}$. S'il en est ainsi, les éléments u_1, u_2 seront racines du polynôme du second degré

$$(X - u_1)(X - u_2) = X^2 - (u_1 + u_2)X + u_1u_2$$

à coefficients complexes. D'après les formules connues on a

$$u_1, u_2 = \frac{u_1 + u_2}{2} \pm \sqrt{\left(\frac{u_1 + u_2}{2}\right)^2 - u_1u_2},$$

si bien que u_1, u_2 sont aussi des nombres complexes. Ainsi, pour le polynôme $f(X)$ à coefficients réels considéré, nous avons trouvé même deux racines complexes.

Soit maintenant

$$f(X) = a_0X^n + a_1X^{n-1} + \ldots + a_{n-1}X + a_n$$

un polynôme de degré n à coefficients complexes arbitraires (on peut poser $a_0 = 1$, mais cela n'a pas d'importance). En remplaçant tous les a_i par leurs conjugués, on obtient un polynôme

$$\bar{f}(X) = \bar{a}_0 X^n + \bar{a}_1X^{n-1} + \ldots + \bar{a}_{n-1}X + \bar{a}_n.$$

Introduisons un polynôme

$$e(X) = f(X)\bar{f}(X) = e_0X^{2n} + e_1X^{2n-1} + \ldots + e_{2n}$$

de degré $2n$ à coefficients

$$e_k = \sum_{i+j=k} a_i\bar{a}_j, \quad k = 0, 1, \ldots, 2n.$$

Puisque l'opération de conjugaison $z \mapsto \bar{z}$ est un automorphisme d'ordre 2 du corps $\mathbb{C}$ (chap. 5, § 1, théorème 1), on a $\bar{e}_k = \sum_{i+j=k} \bar{a}_i a_j = e_k$, ce qui signifie que $e_k \in \mathbb{R}$. En vertu de ce qui a

été démontré, le polynôme $e(X)$ à coefficients réels admet au moins une racine complexe c:

$$f(c)\cdot\bar{f}(c) = e(c) = 0.$$

On en déduit que, ou bien $f(c) = 0$ et le théorème est donc démontré, ou bien $\bar{f}(c) = 0$, c'est-à-dire $\bar{a}_0 c^n + \bar{a}_1 c^{n-1} + \ldots + \bar{a}_{n-1} c + \bar{a}_n = 0$. En appliquant aux deux membres de cette égalité l'automorphisme de conjugaison complexe, on obtient $a_0 \bar{c}_1^n + a_1 \bar{c}^{n-1} + \ldots + a_{n-1}\bar{c} + a_n = 0$, c'est-à-dire $f(\bar{c}) = 0$. ■

La propriété du corps $\mathbb{C}$ d'être algébriquement clos (ainsi que le fait d'existence d'un corps de décomposition des polynômes) est avantageusement utilisée pour la résolution de différents problèmes.

Exemple. — Soient $S_0(f)$ l'ensemble de toutes les racines différentes du polynôme $f \in \mathbb{C}[X]$, et $S_1(f)$ l'ensemble de toutes ses « unités »: $d \in S_1(f) \Leftrightarrow f(d) = 1$. Soient maintenant f, g deux polynômes quelconques de $\mathbb{C}[X]$. Démontrer que

$$S_0(f) = S_0(g),\ S_1(f) = S_1(g) \Rightarrow f(X) = g(X).$$

Puisqu'on a évidemment $S_0(f) \cap S_1(f) = \varnothing$, il suffit, étant donnés les résultats du § 1, de démontrer que $|S_0(f) \cup S_1(f)| \geqslant n + 1$, où $n = \deg f$. D après le théorème 1

$$f(X) = a_0 \prod_{i=1}^{\nu} (X - c_i)^{s_i},\quad f(X) - 1 = a_0 \prod_{j=1}^{\mu} (X - d_j)^{t_j},\quad c_i,\ d_j \in \mathbb{C},$$

où

$$\sum s_i = n = \sum t_j,\quad \nu + \mu = |S_0(f) \cup S_1(f)|.$$

Conformément au théorème 5 du § 1, on a

$$f(X)' = (f(X) - 1)' = \prod_{i=1}^{\nu} (X - c_i)^{s_i - 1} \cdot \prod_{j=1}^{\mu} (X - d_j)^{t_j - 1} \cdot h(X),$$

si bien que $(n - \nu) + (n - \mu) = \sum (s_i - 1) + \sum (t_j - 1) \leqslant \deg f(X)' = n - 1$. Donc,

$$\nu + \mu \geqslant n + 1.$$

§ 4. Polynômes à coefficients réels

1. Décomposition en facteurs irréductibles dans $\mathbb{R}[X]$.— Le théorème 1 du § 3 entraîne que tout polynôme f de degré n peut s'écrire dans $\mathbb{C}[X]$ d'une manière et d'une seule (à l'ordre des facteurs près) sous la forme

$$f(X) = a(X - c_1)(X - c_2) \ldots (X - c_n),$$

où $a \neq 0$, $c_1, \ldots, c_n$ sont des nombres complexes. Soient maintenant $f(X) = X^n + a_1X^{n-1} + \ldots + a_{n-1}X + a_n$ un polynôme unitaire à coefficients réels $a_1, \ldots, a_n$, et c une racine complexe quelconque de ce polynôme: $c = u + iv$, $v \neq 0$. En appliquant à la relation $f(c) = 0$ l'automorphisme de conjugaison complexe, comme nous l'avons fait lors de la démonstration du théorème 1 du § 3, on obtient aussi $f(\bar{c}) = 0$, puisque $\bar{a}_i = a_i$. Par conséquent, le polynôme $f(X)$ est divisible par un polynôme du second degré

$$g(X) = (X - c)(X - \bar{c}) = X^2 - (c + \bar{c})X + c\bar{c} =$$

$$= X^2 - 2uX + (u^2 + v^2)$$

à discriminant négatif $D(g) = 4u^2 - 4(u^2 + v^2) = -4v^2 < 0$. La condition $D(g) < 0$ est nécessaire et suffisante pour que le polynôme du second degré $g \in \mathbb{R}[X]$ soit irréductible sur $\mathbb{R}$.

Si k est la multiplicité de la racine c du polynôme $f(X)$, et $l \leqslant k$ la multiplicité de la racine c, alors le polynôme $f(X)$ est divisible par la puissance l-ième du polynôme $g(X)$:

$$f(X) = g(X)^l q(X).$$

Le quotient $q(X)$ de deux polynômes de $\mathbb{R}[X]$ est encore le polynôme de $\mathbb{R}[X]$. Si $k > l$, l'élément $c \in \mathbb{C}$ est la racine de $q(X)$ de multiplicité $k - l$, alors que $\bar{c}$ ne l'est pas. Or, nous avons vu qu'il n'en est pas ainsi. Il s'ensuit que $k = l$ (la supposition $l \geqslant k$ est considérée de façon analogue), c'est-à-dire que les racines complexes de tout polynôme de $\mathbb{R}[X]$ sont deux à deux conjuguées. Nous arrivons à la conclusion que les éléments de l'anneau factoriel $\mathbb{R}[X]$ vérifient l'assertion suivante:

THÉORÈME 1.— *Tout polynôme unitaire $f \in \mathbb{R}[X]$ de degré n admet une décomposition et une seule (à l'ordre des facteurs près) en un produit de $m \leqslant n$ polynômes linéaires $X - c_i$ qui correspondent à ses racines réelles $c_1, \ldots, c_m$, et de $(n - m)/2$ polynômes du second degré irréductibles sur $\mathbb{R}$ et correspondant aux couples des racines complexes conjuguées.* ■

REMARQUES. — 1) Un polynôme irréductible dans $\mathbb{R}[X]$ est soit linéaire, soit du second degré et à discriminant négatif.

2) Avec les notations du théorème 1, on a la relation

$$D(f) = (-1)^{\frac{n-m}{2}} |D(f)|,$$

c'est-à-dire le signe du discriminant est déterminé par le nombre de couples des racines complexes conjuguées. Cette relation est obtenue soit directement de la définition du discriminant, soit à l'aide de la formule contenue dans l'exercice 5 du § 2.

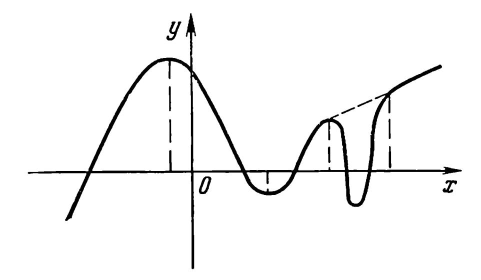
3) Les fractions rationnelles simples ont dans le corps $\mathbb{R}[X]$ la forme (9) (voir chap. 5, § 4).

2. Problème de localisation des racines d'un polynôme.— Considérons le polynôme $f \in \mathbb{R}[X]$ comme une fonction à valeurs réelles $x \mapsto f(x)$ de l'argument réel x, en représentant celle-ci par un graphe du plan rapporté à un repère orthonormé xOy. Aux racines réelles du polynôme $f(X)$ (ou aux zéros de la fonction $f(x)$) correspondent les abscisses des points d'intersection du graphe avec l'axe des x.

La première question importante qu'on a généralement à traiter en pratique concerne le problème d'encadrement des racines réelles, c'est-à-dire la recherche de l'intervalle $a < x < b$ qui comprend toutes les racines réelles d'un polynôme donné f. A proprement parler, nous savons déjà du lemme 1 du § 3 que pour $|x| > \frac{A}{|a_0|} + 1$ (a_0 est le coefficient dominant, $A = \max\{|a_1|, \ldots, |a_n|\}$) la fonction $f(x)$ ne s'annule pas même dans le cas du plan $\mathbb{C}$. Les encadrements plus précis des racines sont indiqués dans les exercices 1 à 4.

Un problème plus général de *localisation* (de *séparation*) *des racines* d'un polynôme consiste à indiquer pour chacune des racines réelles un intervalle à l'intérieur duquel n'est comprise que cette racine. Une solution satisfaisante bien qu'assez encombrante de ce problème a été obtenue pour la première fois par Sturm en 1829. Nous nous contenterons de démontrer des résultats plus particuliers, ayant en vue que la résolution complète du problème de localisation des racines (surtout si l'on tient compte de toutes les racines, y compris les racines complexes quand il s'agit non pas des intervalles mais des domaines sur le plan $\mathbb{C}$) présente des difficultés considérables et que sa simplification pour des classes spéciales de polynômes fait l'objet des préoccupations des spécialistes. Nous laissons de côté les méthodes de calcul d'une « racine localisée » avec un degré de précision donné. Le calcul mathématique moderne dispose dans ce domaine d'un large arsénal de moyens. Une étude détaillée de ces moyens sortirait nettement du cadre de cet ouvrage.

Heureusement, dans beaucoup de cas on peut se contenter d'un aspect qualitatif approché de la disposition des racines. On obtient une information substantielle en construisant le graphe de la fonction $x \mapsto f(x)$ dont les valeurs sont calculées (à l'aide du schéma de Horner, par exemple) ne serait-ce qu'aux points à valeurs entières de l'axe des x. Il faut s'attendre que les racines de l'équation



algébrique $f(x) = 0$ se situent entre les points (ou aux points) où la fonction présente ses maximums et minimums, ceux-ci étant à leur tour racines de l'équation algébrique $f'(x) = 0$ de degré inférieur. L'examen du graphe permet en tout cas de minorer le nombre de racines réelles positives et négatives, à savoir, obtenir des estimations et non pas des valeurs exactes, parce qu'en général nous pouvons ne pas tenir compte des variations de la fonction $x \mapsto f(x)$ dans des intervalles étroits quelconques.

Il est remarquable que les majorations pour les mêmes valeurs sont obtenues à partir des considérations bien simples qui ont été présentées encore par Descartes en 1637. Introduisons la définition suivante:

DÉFINITION. — *Soient*

$$a_0, a_{i_1}, a_{i_2}, \ldots, a_{i_q} \quad (0 < i_1 < i_2 < \ldots < i_q \leqslant n) \qquad (1)$$

tous les coefficients non nuls du polynôme $f(X) = a_0X^n + a_1X^{n-1} + \ldots \ldots \in \mathbb{R}[X]$ *écrits dans cet ordre. Si* $a_{i_k}a_{i_{k+1}} < 0$, *on dit que sur le* $(k+1)$*-ième terme il y a changement de signe. Le nombre total de changements de signe dans la suite* (1) *est désigné par le symbole* $L(f)$.

Il est clair qu'on a toujours $0 \leqslant L(f) \leqslant \deg f$ et qu'en outre $L(-f) = L(f)$. Remarquons aussi que $L(f) = L(aX^k + a_{i_1}X^{n-i_1} + \ldots)$, où l'exposant k satisfait à la condition unique $k > n - i_1$, et $aa_0 > 0$. Si $L(f) = 0$, il est évident que f ne possède pas de racines positives. D'autre part, f peut ne pas posséder des racines positives aussi dans le cas où $L(f) = \deg f$. Un exemple: $f(X) = X^2 - X + 1$. Néanmoins, le symbole $L(f)$ a, comme nous le verrons plus loin, un rapport direct au nombre de racines positives du polynôme f.

LEMME. — *Si* $c > 0$, *alors* $L((X - c)f) = L(f) + 1 + 2s$, *où* $s \in \mathbb{Z}$, $s \geqslant 0$.

DÉMONSTRATION.— On suppose évidemment que $f \neq 0$, si bien que le symbole $L(f)$ a un sens. Si $\deg f = 0$, alors $L(f) = 0$ et le lemme est vrai, avec $s = 0$. En raisonnant par récurrence sur $\deg f$, supposons le lemme démontré pour tous les polynômes de degré $<n$. Soient $\deg f = n$ et

$$f = a_0X^n + a_kX^{n-k} + \ldots + a_{n-1}X + a_n,$$

où a_k est, après a_0, le premier coefficient non nul, s'il existe ($k \geqslant 1$). Puisque $L(-f) = L(f)$, on peut poser, sans restreindre la généralité, $a_0 > 0$. Soit

$$g(X) = a_kX^{n-k} + \ldots + a_{n-1}X + a_n.$$

Il est clair que

$$L(f) = L(g) + \varepsilon, \tag{2}$$

où

$$\varepsilon = \frac{1}{2}\left(1 - \frac{a_0a_k}{|a_0a_k|}\right) = 0 \text{ ou } 1.$$

On suppose de nouveau que $g \neq 0$, sinon la démonstration pour f est évidente. Posons encore pour la suite que

$$(X - c)\, g(X) = a_kX^{n+1-k} + h(X)$$

(remarquons que $g \neq 0 \Rightarrow h \neq 0$).

Par hypothèse de récurrence et en raison de l'égalité (2) on a

$$L((X - c)\, g(X)) = L(g) + 1 + 2t = L(f) + 1 - \varepsilon + 2t. \tag{3}$$

On a aussi

$$(X - c) f = a_0X^n (X - c) + (X - c)\, g =$$
$$= a_0X^{n+1} - a_0cX^n + a_kX^{n+1-k} + h(X).$$

Si $k > 1$, il est évident que $L((X - c) f) = 2 - \varepsilon + L((X - c)\, g)$, car $c > 0$ (ici $2 - \varepsilon$ est le nombre de changements de signe dans la suite a_0, $-a_0c$, a_k). En tenant compte de (3), on obtient

$$L((X - c) f) = L(f) + 1 + 2s, \quad \text{où } s = t + 1 - \varepsilon \geqslant 0.$$

Il ne reste qu'à considérer le cas où $k = 1$:

$$(X - c) f = a_0X^{n+1} + (a_1 - a_0c)\, X^n + h(X).$$

Si a_1 et $a_1 - a_0c$ sont de même signe, on a

$$L((a_1 - a_0c)\, X^n + h(X)) = L((X - c)\, g)$$

et

$$L((X - c) f) = \varepsilon + L((X - c)\, g) = L(f) + 1 + 2s,\ s = t.$$

Lorsque a_1 et $a_1 - a_0c$ sont de signes contraires, ce qui n'est possible que pour $a_1 > 0$ et $\varepsilon = 0$, on a

$$L((a_1 - a_0c)\, X^n + h(X)) = L((X - c)\, g) \pm 1 =$$
$$= L(f) + 1 + 2t \pm 1$$

et

$$L((X-c)f) = 1 + L((a_1 - a_0c)X^n + h(X)) =$$
$$= L(f) + 1 + 2s,$$

où $s = t$ ou $t + 1$. Enfin, dans le cas où $a_1 - a_0c = 0$, ce qui ne peut avoir lieu, comme précédemment, que si $a_1 > 0$ et $\varepsilon = 0$, on a

$$L((X-c)f) = L(a_0X^{n+1} + h(X)) = L(a_1X^n + h(X)) =$$
$$= L((X-c)g) = L(f) + 1 + 2s, \quad s = t. \quad ■$$

En partant du lemme démontré on obtient facilement la *règle des signes due à Descartes*:

THÉORÈME 2. — *Le nombre de racines positives d'un polynôme f à coefficients réels est égal ou inférieur d'un nombre pair à $L(f)$.*

DÉMONSTRATION.— Soient $c_1, c_2, \ldots, c_m$ les racines positives (non nécessairement distinctes) du polynôme $f(X) = a_0X^n + \ldots$ $\ldots + a_{n-\nu}X^\nu$, où, par hypothèse, $a_0 > 0$ et $a_{n-\nu}$ est le dernier coefficient non nul. En utilisant la forme de la décomposition canonique du polynôme (théorème 1) on peut écrire

$$f(X) = (X-c_1)\ldots(X-c_m)\,g(X), \tag{4}$$

où $g(X) = a_0X^{n-m} + \ldots + bX^\nu, a_0 > 0, b > 0\ (\nu \geqslant 0)$. Puisque a_0 et b sont de même signe, $L(g) = 2t$ est un nombre pair. En tenant compte du lemme et de la décomposition (4), on obtient une suite d'égalités

$$L((X-c_1)g) = 1 + 2(s_1 + t),$$
$$L((X-c_2)(X-c_1)g) = 1 + 2(s_1 + t) + 1 + 2s_2 =$$
$$= 2 + 2(s_1 + s_2 + t),$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$L(f) = m + 2(s_1 + s_1 + \ldots + s_m + t).$$

La dernière de celles-ci exprime justement l'assertion énoncée dans le théorème. ■

Ainsi donc, on a toujours $m \leqslant L(f)$. Maintenant, nous allons examiner un cas d'importance pratique, où l'on sait *a priori*, à partir des considérations quelconques, que toutes les racines d'un polynôme f sont réelles. Ceci étant, on peut énoncer le théorème qui apporte une précision à celui de Descartes.

THÉORÈME 3. — *Si toutes les racines d'un polynôme f sont réelles, le nombre $m(f) = m$ de ses racines positives satisfait, compte tenu de leur multiplicité, à l'égalité $m(f) = L(f)$.*

DÉMONSTRATION. — Le théorème 3 pourrait se déduire assez facilement du théorème 2, mais nous allons donner ci-dessous la

démonstration indépendante qui, elle aussi, est également simple (et d'ailleurs instructive).

D'après le théorème de Rolle (ou théorème de la moyenne) bien connu en Analyse, il existe, entre les racines a' et b' de notre polynôme $f(X)$, un nombre $c \in \mathbb{R}$, $a' < c < b'$, tel que $f'(c) = 0$. On en déduit que toutes les racines du polynôme dérivé $f'(X)$ sont réelles et que $m(f') = m(f)$ ou $m(f) - 1$. En effet, soient $c_1 < c_2 < \ldots$ $\ldots < c_r$ les racines du polynôme f d'ordres de multiplicité respectifs $n_1, n_2, \ldots, n_r$, si bien que $n_1 + n_2 + \ldots + n_r = \deg f = n$. En vertu du théorème 5 du § 1, le polynôme dérivé f' possède les racines $c_1, c_2, \ldots, c_r$ d'ordres de multiplicité $n_1 - 1, n_2 - 1, \ldots$ $\ldots, n_r - 1$ et, d'après le théorème de Rolle, encore au moins une racine dans chaque intervalle entre ces racines, à savoir: $c'_1, c'_2, \ldots$ $\ldots, c'_{r-1}$. On obtient au total $(n_1 - 1) + \ldots + (n_r - 1) +$ $+ r - 1 = n - 1$ racines réelles. Puisque $\deg f' = n - 1$, le polynôme dérivé f' n'admet pas d'autres racines. Soient ensuite $c_{l-1} < 0$, et $c_l, \ldots, c_r$ toutes les racines positives d'ordres de multiplicité $n_l, \ldots, n_r$: $n_l + \ldots + n_r = m = m(f)$. Parmi les racines positives du polynôme dérivé $f'(X)$ seront les racines $c_l, \ldots$ $\ldots, c_r$ d'ordres de multiplicité $n_l - 1, \ldots, n_r - 1$, les racines $c'_l, \ldots, c'_{r-1}$ et peut-être encore la racine c'_{l-1}, de sorte que leur nombre total sera $m(f') = m(f) - 1$ ou $m(f)$, comme cela a été énoncé. L'expression analytique de ce fait est donnée par la formule presque tautologique

$$m(f) = m(f') + \varepsilon, \quad \varepsilon = \frac{1}{2}\left(1 - (-1)^{m(f)+m(f')}\right). \tag{5}$$

Observons encore que si

$$f(X) = a_0 X^n + \ldots + a_{n-\nu} X^{\nu}, \tag{6}$$

où $a_{n-\nu}$ est le dernier coefficient non nul, on a, conformément à l'écriture (4), $a_{n-\nu} = (-1)^m c_1 c_2 \ldots c_m b$, où $c_k > 0$ et $b > 0$. En d'autres termes

$$(-1)^{m(f)} a_{n-\nu} > 0. \tag{7}$$

En raisonnant maintenant par récurrence sur $n = \deg f$, supposons le théorème démontré pour tous les polynômes de degré $< n$. Si $\nu > 0$ dans (6), c'est-à-dire $a_n = 0$, alors $f(X) = X \cdot f_1(X)$ et $m(f) = m(f_1) = L(f_1) = L(f)$ (on a par récurrence $m(f_1) =$ $= L(f_1)$). Il reste à considérer le cas de $a_n \neq 0$. Soit

$$f'(X) = na_0 X^{n-1} + \ldots + \mu a_{n-\mu} X^{\mu-1}, \quad a_{n-\mu} \neq 0.$$

Alors,

$$L(f) = L(f') + \delta, \quad \delta = \frac{1}{2}\left(1 - \frac{a_n a_{n-\mu}}{|a_n a_{n-\mu}|}\right) = 0 \text{ ou } 1.$$

Mais nous savons (voir (7)) que $(-1)^{m(f)} a_n > 0$ et $(-1)^{m(f')} a_{n-\mu} >$ > 0. Par conséquent, $\delta = \frac{1}{2}(1 - (-1)^{m(f)+m(f')})$ et donc $\delta = \varepsilon$.

Puisque, par hypothèse de récurrence, $L(f') = m(f')$, on a finalement $L(f) = m(f') + \varepsilon$ ou, en comparant avec (5), $m(f) = L(f)$. ■

COROLLAIRE (cas particulier du théorème de Budan-Fourier). — *Si toutes les racines d'un polynôme f sont réelles, le nombre de ses racines comprises dans l'intervalle (a, b) est égal à $L(f_a) - L(f_b)$, où*

$$f_a(X) = f(X+a) = \sum_{0 \leqslant k \leqslant n} \frac{f^{(k)}(a)}{k!} X^k,$$

$$f_b(X) = f(X+b) = \sum_{0 \leqslant k \leqslant n} \frac{f^{(k)}(b)}{k!} X^k$$

sont les développements en série de Taylor (voir exercice 3).

DÉMONSTRATION. — Le nombre $m(f_a)$ de racines positives du polynôme f_a est, par définition, égal au nombre de racines du polynôme donné f, supérieures à a. La même remarque est valable pour f_b. Par conséquent, le nombre de racines du polynôme f, comprises entre a et b $(a < b)$ est égal à la différence $m(f_a) - m(f_b)$ qui s'exprime, suivant le théorème 2, sous la forme $L(f_a) - L(f_b)$. ■

3. Polynômes stables.— On dit qu'un polynôme unitaire

$$f(X) = X^n + a_1X^{n-1} + \ldots + a_{n-1}X + a_n$$

à coefficients réels est *stable* si toutes ses racines se situent dans le demi-plan à gauche:

$$f(\lambda) = 0,\ \lambda = \alpha + i\beta \Rightarrow \alpha < 0$$

(voir fig. 18). Cette terminologie a son origine dans la théorie des équations différentielles. Les critères de comportement asymptotiquement stable d'un

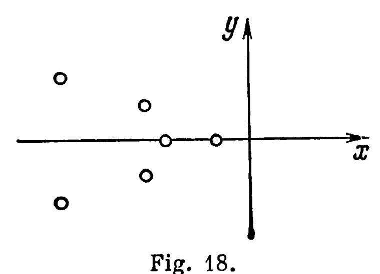


Fig. 18.

système physique (et, dans un sens plus large, d'un système mécanique, technique ou économique) dans le voisinage de la position d'équilibre, élaborés par cette théorie, exigent que

$$\lim_{t \to +\infty} e^{\lambda t} = 0, \tag{8}$$

où λ est une racine arbitraire du polynôme f associé à une équation différentielle d'ordre n à coefficients constants. Puisque, d'après la formule d'Euler (voir chap. 5, § 1, formule (15)) $e^{\lambda t} = e^{\alpha t}e^{i\beta t} = e^{\alpha t}(\cos \beta t + i \sin \beta t)$, le terme dominant est $e^{\alpha t}$, de sorte que la condition (8) est équivalente à l'inégalité $\alpha < 0$.

Il se pose un problème particulier de localisation, *problème de Routh et*

Hurwitz (*) lorsqu'il s'agit de déterminer directement, d'après les coefficients d'un polynôme f, si ce polynôme est stable. Ce problème algébrique a été résolu encore en 1895. Le critère de Routh et Hurwitz dit: *un polynôme f est stable si, et seulement si, sont satisfaites les inégalités*

$$\Gamma_1 > 0,\ \Gamma_2 > 0,\ \ldots,\ \Gamma_n > 0, \tag{9}$$

où

$$\Gamma_k = \begin{vmatrix} a_1 & 1 & 0 & 0 & 0 & 0 & \ldots & 0 \\ a_3 & a_2 & a_1 & 1 & 0 & 0 & \ldots & 0 \\ a_5 & a_4 & a_3 & a_2 & a_1 & 1 & \ldots & 0 \\ a_7 & a_6 & a_5 & a_4 & a_3 & a_2 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{2k-1} & a_{2k-2} & a_{2k-3} & a_{2k-4} & a_{2k-5} & a_{2k-6} & \ldots & a_k \end{vmatrix}$$

(on suppose que $a_s = 0$ pour $s > n$).

Sans tenter de démontrer le théorème de Routh et Hurwitz (cela se fait dans d'autres cours), il est à remarquer que l'élégance de son énoncé est due entièrement à la théorie des déterminants. En vertu du théorème 1, lorsque les conditions (9) sont satisfaites, le polynôme $f(X)$ se représente par un produit des facteurs de la forme $X + u$, $X^2 + vX + w$, avec $u > 0$, $v > 0$, $w > 0$, ce qui signifie que tous les coefficients du polynôme stable $f(X)$ sont positifs:

$$a_1 > 0,\ a_2 > 0,\ \ldots,\ a_n > 0. \tag{10}$$

Ainsi, les conditions (10) sont nécessaires pour que le polynôme $f(X)$ soit stable. N'étant pas, en général, suffisantes, ces conditions permettent quand même de réduire à peu près de moitié le nombre d'inégalités des déterminants (9). C'est très commode parce que le calcul des déterminants est bien fastidieux.

EXEMPLE. — Pour $n = 2$, le système d'inégalités $\Gamma_1 > 0$, $\Gamma_2 > 0$ est équivalent à un système plus simple: $a_1 > 0$, $a_2 > 0$, ce qui d'ailleurs découle des formules donnant les racines de l'équation du second degré.

Pour $n = 3$, tout se ramène aux inégalités $a_1 > 0$, $a_2 > 0$, $a_3 > 0$, $a_1a_2 > a_3$, puisque $\Gamma_3 = a_3(a_1a_2 - a_3)$.

Signalons enfin que le critère de Routh et Hurwitz ne résout pas toutes les questions liées à la stabilité, parce que, dans la pratique, il s'agit des polynômes et des équations différentielles dont les coefficients dépendent d'un paramètre. C'est en termes du paramètre lui-même que doivent être énoncées les conditions de stabilité, ce qui pose un problème de nature tout à fait différente.

EXERCICES

1. Soit $f(X) = a_0X^n + a_1X^{n-1} + \ldots + a_n$ un polynôme à coefficients réels de degré n. Montrer que la connaissance des majorants des racines positives des polynômes $f(X)$, $X^nf\left(\frac{1}{X}\right)$, $f(-X)$, $X^nf\left(-\frac{1}{X}\right)$ donne les minorants et les majorants des racines tant positives que négatives du polynôme $f(X)$.

2. Soient, avec les notations de l'exercice 1, $a_0 > 0$, m le plus petit indice pour lequel $a_m < 0$, B le maximum des valeurs absolues des coefficients négatifs. Montrer que

$$c \leqslant 1 + \sqrt[m]{B/a_0}$$

pour toute racine réelle positive du polynôme $f(X)$. (Indication. Pour $x > 1$, partir de l'estimation

$$f(x) \geqslant a_0x^n - B\,\frac{x^{n-m+1} - 1}{x-1} > \frac{x^{n-m+1}}{x-1}\,[a_0x^{m-1}(x-1) - B].)$$

(*) En réalité, ce problème a été posé beaucoup plus tôt (1868) par le physicien anglais Maxwell et résolu pour de faibles degrés par l'ingénieur russe I. A. Vychnégradsky qui s'occupait du problème de stabilité des régulateurs (1876).

3 (*formule de Taylor*). Soit P un corps commutatif de caractéristique nulle, $a \in P$. Montrer que la formule

$$f(X)=f(a)+\frac{f'(a)}{1!}(X-a)+\frac{f''(a)}{2!}(X-a)^2+\ldots+\frac{f^{(n)}(a)}{n!}(X-a)^n$$

est valable pour tout polynôme $f \in P[X]$ de degré n. (Indication. Dériver k fois l'expression formelle $f(X) = \sum b_i (X - a)^i$ et poser $X = a$.)

4. Montrer que, si $f(a) > 0$, $f'(a) > 0, \ldots, f^{(n)}(a) > 0$ pour un polynôme de degré n à coefficients réels $f(X)$ et à coefficient dominant positif a_0, alors $f(c) = 0$, $c > 0 \Rightarrow c < a$. (Indication. Se servir de l'exercice 3.)

5. En utilisant la règle des signes de Descartes, déterminer le signe du discriminant des polynômes $X^5 - X^2 + 1$, $X^3 - 6X - 9$ (voir remarque en fin du nº 1).

6. Les polynômes $X^5 - X - 1$ et $X^3 + aX + b \in \mathbb{Q}[X]$ peuvent-ils avoir des racines complexes communes? Rappelons (voir § 1, exercice 11) que le polynôme $X^5 - X - 1$ est irréductible sur $\mathbb{Q}$.

7. Montrer que les racines du polynôme $f(X) = X^5 + uX^4 + vX^3 + w \in \mathbb{R}[X]$ à terme constant $w \neq 0$ ne peuvent pas être toutes réelles (Indication. Il est commode de passer au polynôme $X^5 f\left(\frac{1}{X}\right)$ et utiliser ensuite les formules (12) du § 1 et (9) du § 2.)

8. Il est clair que si un polynôme à coefficients entiers $f(X) = a_0X^n + \ldots + a_n$ possède une racine $c \in \mathbb{Z}$, alors c divise le terme constant $a_n = f(0)$: $f(c) = 0 \Rightarrow a_n = c(-a_0c^{n-1} - \ldots - a_{n-2}c - a_{n-1})$. Montrer qu'en même temps $c - 1$ divise $f(1) = \sum a_i$, et $c + 1$ divise $f(-1) = \sum (-1)^i a_i$. (Indication. $f(X) = (X - c)\, g(X) \Rightarrow g(X) \in \mathbb{Z}[X]$.) Appliquer ces considérations à la recherche des racines entières du polynôme $X^4 + X^3 - X^2 + 40X - 100$ (réponse: $c = 2$).

9. S'assurer que

$$f(X) = X^n + a_1X^{n-1} + \ldots + a_n \in \mathbb{Z}[X],\ f(c) = 0,\ c \in \mathbb{Q} \Rightarrow c \in \mathbb{Z}.$$

(Indication. Si $c = a/b$ est une fraction irréductible, on a $a^n/b = -a_1a^{n-1} - a_2a^{n-2}b - \ldots - a_nb^{n-1}$.) Que peut-on dire des racines rationnelles d'un polynôme à coefficients entiers dont le coefficient dominant est $a_0 \neq 1$?

10. Montrer que tout polynôme $f(X)$, avec $f(x) \geqslant 0$ pour tous les $x \in \mathbb{R}$, peut être représenté sous la forme

$$f(X) = g(X)^2 + h(X)^2,$$

où $g, h \in \mathbb{R}[X]$. (Indication. En se servant du théorème 1, décomposer $f(X)$ en facteurs de la forme $(X + a)^2 + b^2$ et utiliser l'identité formelle

$$(p^2 + q^2)(r^2 + s^2) = (pr + qs)^2 + (ps - qr)^2,$$

qui résulte de la relation

$$|p + iq|^2\, |r + is|^2 = |(p + iq)(r + is)|^2.)$$

11. Etablir de manière indépendante le critère de stabilité des polynômes de degrés 3 et 4. Pour $n = 4$, l'écrire sous la forme des inégalités $a_1 > 0$, $a_4 > 0$, $a_1a_2 > a_3$, $a_3(a_1a_2 - a_3) > a_1^2a_4$. (Indication. $f(X) = X^3 + aX^2 + bX + c = (X^2 + \alpha X + \beta)(X + \theta)$, où $a = \alpha + \theta$, $b = \beta + \alpha\theta$, $c = \beta\theta$, avec $\alpha, \beta, \theta \in \mathbb{R}$. La stabilité de $f(X)$ est équivalente à celle du couple de polynômes $X^2 + \alpha X + \beta$, $X + \theta$, c'est-à-dire à la réalisation des inégalités $\alpha > 0$, $\beta > 0$, $\theta > 0$. On vérifie facilement que ce système est équivalent au système d'inégalités $a > 0$, $b > 0$, $c > 0$, $ab - c > 0$. Appliquer des considérations analogues au polynôme à coefficients réels de quatrième degré.)

PARTIE II

GROUPES. ANNEAUX. MODULES

Le contenu de la deuxième partie peut être qualifié comme une suite assez sérieuse, mais, espérons-le, pas trop abstraite de la première partie. Les nouvelles notions introduites sont relativement peu nombreuses. Le lecteur rencontrera ses vieilles connaissances du chapitre 4, qui l'introduiront dans le domaine des notions sensiblement plus riches et profondes. Il est recommandé d'attacher la plus grande attention à l'étude des exemples qui occupent un bon quart de tout le texte (notamment, il est naturel de ranger le matériel du chapitre 7, § 1, et du chapitre 8, § 3, dans la catégorie des exemples). En outre, les exemples sont choisis de manière à établir un pont entre l'algèbre et les autres branches de la mathématique. Si, après l'étude de cette deuxième partie, le lecteur ressent un plus fort sentiment de l'unité des mathématiques, le but poursuivi par l'auteur sera atteint.

BIBLIOGRAPHIE

1. Atiyah M., Macdonald I., *Introduction to Commutative Algebra*, London, 1969.
2. Bhagavantam S., Venkatarayudu T., *Theory of groups and its application to physical problems*, Andhra university, Waltair, 1951.
3. Birkhoff G., Bartee T., *Modern Applied Algebra*, N.Y., 1970.
4. Borevich Z. Y., Shafarevich Y. R., *Number theory*. Volume in the pure and appl. Math. Series Acad. Press, New York, 1966.
5. Bourbaki N., *Eléments de mathématique*, Livre II, Algèbre, 1942-1958.
6. Cohn P., *Universal Algebra*, N.Y., 1965.
7. Dieudonné J., Carrell J., *Invariant Theory, Old and New*, N.Y., 1971.
8. Faith C., *Algebra: rings, modules and categories*, Springer-Verlag, Berlin, Heidelberg, N.Y. 1973.
9. Hall M., *The theory of groups*, N.Y., 1959.
10. Herstein I., *Noncommutative Rings*, N.Y., 1968.
11. Jacobson N., *Lie Algebras*, N.Y., 1962.
12. Kirillov A. A., *Eléments de la théorie des représentations*, Editions Mir, Moscou, 1974.

13. Kurosh A. G., *Algèbre générale*, Dunod, Paris, 1967.
14. Mumford D., *Geometric Invariant Theory*, Berlin-Ouest, 1965.
15. Naïmark M., Stern A., *Théorie des représentations des groupes*, Editions Mir, Moscou, 1979.
16. Pontriaguine L. S., trad. anglaise, Princeton, 1939; trad. allemande Leipzig, 1957; trad. espagnole, Moscou, 1978 (*Les groupes continus*, en russe).
17. Serre J.-P., *Cours d'arithmétique*, Paris, 1970.
18. Serre J.-P., *Représentations linéaires des groupes finis*, Paris, 1967.
19. Weyl H., *The classical groups, their invariants and representations*, Princeton, 1939; 2e édition, 1946.

CHAPITRE 7

GROUPES

Le présent chapitre se propose de développer la notion de groupe introduite au chapitre 4. L'accent est mis en premier lieu non pas sur les groupes abstraits qui font l'objet de nombreux ouvrages, mais sur l'étude de différentes « opérations » naturelles des groupes. Ce sont justement les réalisations concrètes des groupes qui ont donné une impulsion décisive au développement de la théorie générale des groupes et lui ont valu la réputation d'un instrument utile pour les études mathématiques.

Sur le fond des exemples particuliers (mais, remarquons-le, importants), il s'impose, avec une insistance accrue, l'idée de considérer les (homo-, épi-, iso-) morphismes des groupes, ainsi que les structures réalisées en théorie des groupes, comme un appareil permettant de ramener l'étude des êtres complexes à celle des êtres plus simples.

§ 1. Groupes classiques de faible dimension

1. Définitions générales.— Le cours d'Algèbre linéaire et de Géométrie nous fournit de nouveaux types de groupes qui méritent d'être étudiés avec plus de détails. L'examen des sous-groupes des groupes de transformations des espaces affines, euclidiens et hermitiens, qui laissent invariant un point fixe (par exemple l'origine des coordonnées) nous conduit à des *groupes* dits *classiques* GL (n), SL (n), O (n), SO (n), U (n), SU (n). Signalons que leur vraie place est parmi les groupes dits de Lie. Il faudrait ajouter encore au moins le groupe symplectique Sp (n), mais nous n'avons point l'intention de décrire ici tous les groupes classiques, cela se fait dans d'autres livres. Dans le cas de petites valeurs de n on parle des groupes classiques de faible dimension. Quant aux groupes GL (n), SL (n), nous les avons déjà rencontrés (voir partie I). Désirant éviter une trop forte dépendance vis-à-vis de la géométrie, rappelons que le choix d'une base orthonormée dans un espace conduit à une défini-

tion matricielle équivalente des groupes orthogonal et unitaire:

$$\mathrm{O}(n) = \{A \in M_n(\mathbb{R}) \mid {}^tA \cdot A = A \cdot {}^tA = E\},$$
$$\mathrm{SO}(n) = \{A \in \mathrm{O}(n) \mid \det A = 1\},$$
$$\mathrm{U}(n) = \{A \in M_n(\mathbb{C}) \mid A^* \cdot A = A \cdot A^* = E\},$$
$$\mathrm{SU}(n) = \{A \in \mathrm{U}(n) \mid \det A = 1\}.$$

Ici, $A^* = {}^t\bar{A}$ est une matrice transposée de $A = (a_{ij})$ dont les coefficients a_{ij} sont remplacés par les nombres conjugués $\bar{a}_{ij}$. Les groupes SL (n), SO (n), SU (n) portent le nom de *groupes spéciaux* (*linéaires*, *orthogonaux* et *unitaires*). En particulier,

$$\mathrm{O}(1) = \{\pm 1\},\quad \mathrm{SO}(1) = \{1\},$$
$$\mathrm{U}(1) = \{e^{i\varphi} \mid 0 \leqslant \varphi < 2\pi\},\quad \mathrm{SU}(1) = \{1\},$$
$$\mathrm{SO}(2) = \left\{ \left\| \begin{matrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{matrix} \right\| 0 \leqslant \varphi < 2\pi \right\} \cong \mathrm{U}(1).$$

L'isomorphisme entre les groupes SO (2) et U (1) est donné par la correspondance naturelle

$$\left\| \begin{matrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{matrix} \right\| \mapsto e^{i\varphi}.$$

La représentation géométrique des nombres complexes $e^{i\varphi}$, $0 \leqslant \varphi < 2\pi$, étant le cercle unité S^1 de $\mathbb{R}^2$, on dit encore que le groupe SO(2) et le cercle S^1 sont topologiquement équivalents. Le sens exact de cette terminologie est expliqué dans le cours de Géométrie.

Une relation remarquable, mais beaucoup moins évidente, existe aussi entre les groupes SU(2) et SO(3), Proposons-nous d'établir au préalable la représentation géométrique du groupe SU(2), qui nous conduira par la suite à celle du groupe SO(3).

2. Paramétrisation des groupes SU (2), SO (3).— D'après le théorème d'Euler bien connu, tout élément du groupe SO(3) des rotations propres de l'espace euclidien à trois dimensions $\mathbb{R}^3$ est une rotation autour d'un certain axe fixe. Par exemple, les matrices

$$B_\varphi = \left\| \begin{matrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{matrix} \right\|, \quad C_\theta = \left\| \begin{matrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{matrix} \right\| \tag{1}$$

correspondent à des rotations autour des axes Oz et Ox respectivement d'angles φ et θ. En utilisant la paramétrisation des rotations par les angles d'Euler φ, θ, ψ $(0 \leqslant \varphi, \psi < 2\pi,\ 0 \leqslant \theta < \pi)$ dont le sens géométrique ne nous intéresse pas pour l'instant, on peut écrire

toute matrice $A \in SO(3)$ sous la forme

$$A = B_\varphi C_\theta B_\psi, \tag{2}$$

où B_φ, C_θ, B_ψ sont les matrices (1) indiquées plus haut. Soit ensuite

$$g = \begin{Vmatrix} \alpha & \beta \\ \gamma & \delta \end{Vmatrix} \in SU(2).$$

On a

$$g^* = {}^t\bar{g} = \begin{Vmatrix} \bar{\alpha} & \bar{\gamma} \\ \bar{\beta} & \bar{\delta} \end{Vmatrix}, \quad g^{-1} = \begin{Vmatrix} \delta & -\beta \\ -\gamma & \alpha \end{Vmatrix}.$$

Puisque $g \in U(2) \Longleftrightarrow g^* = g^{-1}$, on a $\delta = \bar{\alpha}$ et $\gamma = -\bar{\beta}$. Ainsi, toute matrice g de SU (2) est de la forme

$$g = \begin{Vmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{Vmatrix}, \quad |\alpha|^2 + |\beta|^2 = 1. \tag{3}$$

Réciproquement, si g est une matrice de la forme (3), il est évident que $g \in SU(2)$. Cela signifie que tout élément du groupe SU(2) est défini de façon unique par un couple de nombres complexes α, β tels que $|\alpha|^2 + |\beta|^2 = 1$. Si l'on pose $\alpha = \alpha_1 + i\alpha_2$, $\beta = \beta_1 + i\beta_2$ avec $\alpha_k, \beta_k \in \mathbb{R}$, $i = \sqrt{-1}$, la condition $|\alpha|^2 + |\beta|^2 = 1$, écrite sous la forme

$$\alpha_1^2 + \alpha_2^2 + \beta_1^2 + \beta_2^2 = 1,$$

permet de dire que le *groupe* SU(2) *est topologiquement équivalent* (*homéomorphe*) *à une sphère* S^3 *de l'espace réel à quatre dimensions* $\mathbb{R}^4$.

Fixons notre attention sur les matrices unitaires

$$b_\varphi = \begin{Vmatrix} e^{i\frac{\varphi}{2}} & 0 \\ 0 & e^{-i\frac{\varphi}{2}} \end{Vmatrix}, \quad c_\theta = \begin{Vmatrix} \cos\frac{\theta}{2} & i\sin\frac{\theta}{2} \\ i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{Vmatrix}. \tag{4}$$

On démontre en Algèbre linéaire (et l'on vérifie directement dans le cas considéré) que pour toute matrice unitaire g de la forme (3) il existe une matrice unitaire u telle que l'on a

$$g = ub_\varphi u^{-1}, \tag{5}$$

avec $\lambda = e^{i\frac{\varphi}{2}}$ déterminé par l'équation du second degré

$$\lambda^2 - 2\alpha_1\lambda + 1 = 0.$$

Remarquons aussi que pour $\alpha\beta \neq 0$ toute matrice (3) peut être mise sous la forme

$$a(\varphi, \theta, \psi) \equiv b_\varphi c_\theta b_\psi = \left\| \begin{matrix} \cos\frac{\theta}{2}\cdot e^{i\frac{\varphi+\psi}{2}} & i\sin\frac{\theta}{2}\cdot e^{i\frac{\varphi-\psi}{2}} \\ i\sin\frac{\theta}{2}\cdot e^{i\frac{\psi-\varphi}{2}} & \cos\frac{\theta}{2}\cdot e^{-i\frac{\varphi+\psi}{2}} \end{matrix} \right\|, \tag{6}$$

où

$$0 \leqslant \varphi < 2\pi,\ 0 \leqslant \theta < \pi,\quad -2\pi \leqslant \psi < 2\pi\ ^{*)}.$$

Il suffit de poser

$$|\alpha| = \cos\frac{\theta}{2},\quad \operatorname{Arg}\alpha = \frac{\varphi+\psi}{2},\qquad |\beta| = \sin\frac{\theta}{2},\ ,\quad \operatorname{Arg}\beta = \frac{\varphi-\psi+\pi}{2},$$

en utilisant le fait que tout nombre complexe z est défini par deux paramètres réels $|z|$ et $\arg z$ ($\operatorname{Arg} z$ est valeur principale de l'argument $\arg z$).

Nous sommes maintenant prêts à aborder la résolution du problème principal de ce paragraphe.

3. Epimorphisme SU (2) → SO (3).— A tout vecteur $x = x_1e_1 + x_2e_2 + x_3e_3$ de l'espace euclidien à trois dimensions $\mathbb{R}^3$, muni de la norme $N(x) = x_1^2 + x_2^2 + x_3^2$, nous faisons correspondre une matrice complexe du deuxième ordre

$$H_x = \left\| \begin{matrix} x_3 & x_1 + ix_2 \\ x_1 - ix_2 & -x_3 \end{matrix} \right\|. \tag{7}$$

L'espace M_2^+ des matrices de la forme (7) est constitué de toutes les matrices hermitiennes de trace nulle (${}^t\overline{H}_x = H_x$, $\operatorname{tr} H_x = 0$), et la correspondance entre les vecteurs $x \in \mathbb{R}^3$ et les matrices $H_x \in M_2^+$ est évidemment biunivoque. En particulier, aux vecteurs de base $e_1, e_2, e_3 \in \mathbb{R}^3$ correspondent les matrices de base $h_k = H_{e_k}$:

$$h_1 = \left\| \begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix} \right\|,\quad h_2 = \left\| \begin{matrix} 0 & i \\ -i & 0 \end{matrix} \right\|,\quad h_3 = \left\| \begin{matrix} 1 & 0 \\ 0 & -1 \end{matrix} \right\|;$$

$$H_x = x_1h_1 + x_2h_2 + x_3h_3,\ M_2^+ = \langle h_1, h_2, h_3 \rangle_{\mathbf{R}}. \tag{8}$$

Remarquons qu'à chaque opérateur linéaire $\Phi^+ : H_x \mapsto H_y$ sur M_2^+ de matrice A dans la base (8), il correspond un opérateur linéaire bien déterminé $\Phi : x \mapsto y$ sur $\mathbb{R}^3$ avec la même matrice A dans la base e_1, e_2, e_3, car $H_{\alpha x} = \alpha H_x$, $H_{x+x'} = H_x + H_{x'}$. Puisqu'aucune autre base n'est utilisée par la suite, il sera commode d'identifier parfois les opérateurs et les matrices qui leur correspondent.

---

*) On verra par la suite que φ, θ, ψ sont les mêmes angles d'Euler. Aux matrices unitaires $\pm g$ on fait correspondre une seule et même rotation dans $\mathbb{R}^3$, et de ce fait, le domaine de variation de ψ se réduit à l'intervalle semi-ouvert $[0,2\pi)$.

Soit maintenant g un élément fixe du groupe SU(2). Considérons l'application

$$\Phi_g^+ : H_x \mapsto g H_x g^{-1}. \quad (9)$$

Puisque les traces de ces matrices coïncident, on a $\operatorname{tr} \Phi_g^+ (H_x) = \operatorname{tr} H_x = 0$. De plus, $g^* = {}^t\bar{g} = g^{-1}$, ce qui implique

$$(gH_xg^{-1})^* = (g^{-1})^* H_x^* g^* = gH_xg^{-1}$$

et, donc,

$$\Phi_g^+ (H_x) \in M_2^+$$

$$\Phi_g^+ (H_x) = \begin{Vmatrix} y_3 & y_1 + iy_2 \\ y_1 - iy_2 & -y_3 \end{Vmatrix} = H_y,$$

où $y = (y_1, y_2, y_3) \in \mathbb{R}^3$. Par suite des égalités (7) et (9) on a

$$\Phi_g^+ (H_{\alpha x + \alpha' x'}) = \alpha \Phi_g^+ (H_x) + \alpha' \Phi_g^+ (H_{x'}),$$

ce qui veut dire que l'application Φ_g^+ (respectivement Φ_g) est un opérateur linéaire sur M_2^+ (respectivement sur $\mathbb{R}^3$).

Montrons que $\Phi_g: \mathbb{R}^3 \to \mathbb{R}^3$ est un opérateur orthogonal. En effet,

$$N (\Phi_g (x)) = N (y) = y_1^2 + y_2^2 + y_3^2 = -\det H_y = -\det \Phi_g^+ (H_x) = -\det gH_xg^{-1} = -\det H_x = x_1^2 + x_2^2 + x_3^2 = N (x),$$

c'est-à-dire Φ_g conserve la norme et, par conséquent, le produit scalaire. Pour le moment, il n'est pas clair si Φ_g change l'orientation de l'espace $\mathbb{R}^3$, ce qui dépend du signe de $\det \Phi_g$. Nous savons seulement que $\det \Phi_g = \pm 1$. Il résulte de la définition que

$$\Phi_{g_1}^+ (\Phi_{g_2}^+ H_x) = g_1 (g_2 H_x g_2^{-1}) g_1^{-1} = (g_1 g_2) H_x (g_1 g_2)^{-1} = \Phi_{g_1 g_2}^+ (H_x),$$

Φ_E^+ étant une matrice unité orthogonale d'ordre 3 pour $E = \begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix} \in \mathrm{SU}(2)$. Cela signifie que la correspondance

$$\Phi : g \mapsto \Phi_g \quad (\text{ou } \Phi^+ : g \mapsto \Phi_g^+)$$

est un homomorphisme de SU(2) dans O(3). Le noyau $\operatorname{Ker} \Phi = \operatorname{Ker} \Phi^+$ se compose de matrices unitaires g pour lesquelles $\Phi_g^+ = \Phi_E^+$. En d'autres termes,

$$\operatorname{Ker} \Phi = \{g \in \mathrm{SU}(2) \mid gH = Hg, \ \forall H \in M_2^+\} = \{g \in \mathrm{SU}(2) \mid gh_j = h_j g, \ j = 1, 2, 3\},$$

où h_1, h_2, h_3 est la base (8) de l'espace M_2^+. Une vérification directe montre que

$$g = \begin{Vmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{Vmatrix}, \ gh_j = h_j g, \ 1 \leqslant j \leqslant 3 \Rightarrow g = \pm E \Rightarrow \operatorname{Ker} \Phi = \{\pm E\}.$$

Examinons maintenant les images des matrices unitaires (4) par l'homomorphisme Φ. Effectuons les calculs pour Φ^+ dans la base (8) :

$$b_\varphi h_1 b_\varphi^{-1} = (\cos\varphi)\, h_1 + (\sin\varphi)\, h_2,$$
$$b_\varphi h_2 b_\varphi^{-1} = (-\sin\varphi)\, h_1 + (\cos\varphi)\, h_2,$$
$$b_\varphi h_3 b_\varphi^{-1} = h_3.$$

Par conséquent (nous passons ici librement de Φ^+ à Φ et des matrices aux opérateurs), $\Phi_{b_\varphi} = B_\varphi$ (voir (1)) est la rotation d'angle φ de l'espace euclidien à trois dimensions $\mathbb{R}^3$ autour de l'axe Ox_3 (ou h_3). En choisissant φ et u de façon à vérifier la relation (5), et compte tenu de ce que Φ est un homomorphisme, on aura

$$\Phi_g = \Phi_u \Phi_{b_\varphi} \Phi_u^{-1} \text{ et } \det \Phi_g = \det \Phi_u \cdot 1 \cdot (\det \Phi_u)^{-1} = 1,$$

ce qui montre qu'en réalité Φ est un homomorphisme de SU(2) dans SO(3).

On vérifie de même que $\Phi_{c_\theta} = C_\theta$ est une rotation d'angle θ autour de l'axe Ox_1, et l'on a maintenant, pour toute matrice $A \in$ $\in$ SO(3):

$$A = B_\varphi C_\theta B_\psi = \Phi_{b_\varphi} \Phi_{c_\theta} \Phi_{b_\psi} = \Phi_{b_\varphi c_\theta b_\psi} = \Phi_{a(\varphi,\ \theta,\ \psi)}.$$

Ainsi donc, l'image Im Φ contient tout le groupe SO(3), et nous avons démontré le théorème suivant :

THÉORÈME 1. — *Le groupe* SO(3) *est une image homomorphe du groupe* SU(2) *par l'homomorphisme* $\Phi : g \mapsto \Phi_g$ *de noyau* Ker $\Phi =$ $= \{\pm E\}$. *Toute rotation de* SO(3) *correspond exactement à deux opérateurs unitaires* g *et* $-g$ *de* SU(2). ■

4. Interprétation géométrique du groupe SO (3). — Du théorème 1 on déduit immédiatement le corollaire suivant:

COROLLAIRE.— *Le groupe* SO (3) *est topologiquement équivalent (homéomorphe) à un espace réel projectif* $\mathbb{R}$ (P^3) *de dimension trois.*

En effet, nous avons vu au nº 2 que les éléments du groupe SU (2) sont en bijection avec les points de la sphère S^3 de l'espace réel à quatre dimensions $\mathbb{R}^4$. Aux opérateurs linéaires $\pm g \in$ SU (2) correspondent des points diamétralement opposés sur S^3, qui se collent par l'homomorphisme Φ. On obtient l'un des modèles de l'espace projectif $\mathbb{R}$ (P^3). ■

En Algèbre linéaire et en Géométrie, l'espace projectif $\mathbb{R}$ (P^n) est défini comme ensemble des droites de l'espace $\mathbb{R}^{n+1}$ passant par l'origine des coordonnées O. Chacune de ces droites coupe la sphère S^3 de rayon unité et de centre O exactement en deux points diamétralement opposés. La donnée de l'un de ces points rétablit de façon univoque la droite correspondante. Or, cela signifie justement que l'espace $\mathbb{R}(P^n)$ peut être défini comme espace quotient de la sphère unité S^n de $\mathbb{R}^{n+1}$ par la relation qui établit l'équivalence des points diamétralement opposés de la sphère S^n. Pour l'instant, nous n'avons pas l'intention de définir la topologie sur $\mathbb{R}(P^n)$.

Nous avons obtenu un résultat assez inattendu. La sphère S^3 et l'espace projectif $\mathbb{R}(P^3)$ possèdent la structure de groupe: SU(2) dans le premier cas et SO(3) dans le deuxième. Toute tentative faite en vue de définir une structure

de groupe continu sur S^2 ou sur $\mathbb{R}(P^2)$ sera vaine (un résultat qui ne se rapporte pas à notre sujet).

D'après le théorème 1 et son corollaire, le groupe SO(3) est « deux fois plus petit » que le groupe SU(2). Vu l'existence de l'épimorphisme SU (2) $\to$ SO (3), il est naturel de se demander s'il existe un monomorphisme SO (3) $\to$ SU(2). Nous verrons au chapitre 8 que la réponse à cette question est négative.

EXERCICES

1. Combler les lacunes dans la démonstration du théorème 1, c'est-à-dire vérifier (sans se référer au cours d'Algèbre linéaire et de Géométrie) toutes les assertions à commencer par l'égalité (2).

2. En utilisant l'interprétation géométrique du groupe SU(2), montrer que

$$(0, 1, 0, 0) * (0, 0, 1, 0) = (0, 0, 0, 1) \neq (0, 0, 1, 0) * (0, 1, 0, 0)$$

(produit des points sur S^3). Les mêmes points (0, 1, 0, 0), (0, 0, 1, 0) considérés dans $\mathbb{R}(P^3)$ sont permutables.

3. Montrer que, si l'on dérive les coefficients des matrices unitaires

$$K_1(t)= \left\| \begin{matrix} \cos\frac{t}{2} & i\sin\frac{t}{2} \\ i\sin\frac{t}{2} & \cos\frac{t}{2} \end{matrix} \right\|, \quad K_2(t)= \left\| \begin{matrix} \cos\frac{t}{2} & -\sin\frac{t}{2} \\ \sin\frac{t}{2} & \cos\frac{t}{2} \end{matrix} \right\|, \quad K_3(t)= \left\| \begin{matrix} e^{i\frac{t}{2}} & 0 \\ 0 & e^{-i\frac{t}{2}} \end{matrix} \right\|$$

par rapport à t et si l'on pose ensuite $t = 0$, on obtient les matrices

$$K_1 = \frac{i}{2}\left\| \begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix} \right\| = \frac{i}{2} h_1, \quad K_2 = \frac{i}{2}\left\| \begin{matrix} 0 & i \\ -i & 0 \end{matrix} \right\| = \frac{i}{2} h_2, \quad K_3 = \frac{i}{2}\left\| \begin{matrix} 1 & 0 \\ 0 & -1 \end{matrix} \right\| = \frac{i}{2} h_3,$$

qui constituent la base de l'espace M_2^- des *matrices hermitiennes gauches*

$$K = \left\| \begin{matrix} ik_3 & -k_2+ik_1 \\ k_2+ik_1 & -ik_3 \end{matrix} \right\|, \quad k_j \in \mathbb{R},$$

ayant une trace nulle: $K^* = -K$, $\operatorname{tr} K = 0$.

§ 2. Opérations des groupes sur les ensembles

1. Homomorphismes $G \to S(\Omega)$.— La théorie des groupes a commencé pour nous au chapitre 4 par des exemples de groupes de transformations, c'est-à-dire des sous-groupes du groupe $S(\Omega)$ de toutes les applications bijectives de l'ensemble Ω sur lui-même. Une telle approche correspond tant à la voie historique de développement de la théorie des groupes qu'à l'importance que les groupes de transformations présentent pour d'autres branches de la mathématique. Bien que la théorie des groupes, dite abstraite, fruit d'une époque plus récente (première moitié du XXe siècle), se soit fort éloignée des groupes de transformations, beaucoup de ses notions portent l'empreinte des temps plus reculés. A savoir, la source de ces notions repose le plus souvent sur l'idée de *réalisation* (*de représentation*) d'un groupe donné G dans $S(\Omega)$, où Ω est un ensemble convenablement choisi. Par réalisation de G dans $S(\Omega)$, il est commode d'en-

tendre tout homomorphisme $\Phi : G \to S(\Omega)$. Si Φ_g est une transformation (un élément) de $S(\Omega)$ correspondant à l'élément $g \in G$, alors $\Phi_e = e_\Omega$ est la transformation identique $\Omega \to \Omega$, et $\Phi_{gh} = \Phi_g \circ \Phi_h$, $g, h \in G$. L'image $\Phi_g(x)$ d'un point (d'un élément) $x \in \Omega$ par la transformation Φ_g est souvent désignée tout simplement par le symbole gx, ce qui permet de parler de l'application $(g, x) \mapsto gx$ du produit cartésien (G, Ω) dans Ω. Il serait plus correct d'écrire $g \circ x$ ou $g * x$ pour éviter toute confusion avec la multiplication dans G, mais dans la plupart des cas cela n'est pas nécessaire. Les propriétés de la transformation Φ_g indiquées plus haut s'écrivent sous la forme

$$\text{(i)} \qquad ex = x, \qquad x \in \Omega;$$

$$\text{(ii)} \qquad (gh)\, x = g\, (hx); \; g, h \in G.$$

Chaque fois qu'il y a l'application $(g, x) \mapsto gx$ du produit cartésien $G \times \Omega$ dans Ω, vérifiant les propriétés (i), (ii), on dit que le groupe *opère* (à gauche) sur l'ensemble Ω et que Ω est un *G-ensemble.* D'autre part, ayant un G-ensemble Ω, nous pouvons, à l'aide de la formule

$$\Phi_g(x) = gx, \quad x \in \Omega,$$

définir pour tout $g \in G$ une application $\Phi_g \colon \Omega \to \Omega$. Ceci étant les propriétés (i), (ii) entraînent que $\Phi : g \mapsto \Phi_g$ est un homomorphisme de G dans $S(\Omega)$. On dit encore (surtout lorsque $|\Omega| < \infty$) qu'à l'opération de G sur Ω est associée la *représentation* (Φ, Ω) *du groupe G dans le groupe de permutations.* Le noyau Ker Φ s'appelle *noyau de l'opération* du groupe G. Si Φ est un monomorphisme (autrement dit, si $gx = x, \forall x \in \Omega \Rightarrow g = e$), on dit que le groupe G opère *effectivement* sur l'ensemble Ω.

Remarque. — Toute opération de G sur Ω induit une opération de G sur $\Omega^k = \Omega \times \ldots \times \Omega$ suivant la règle évidente $g \cdot (x_1, \ldots, x_k) = (gx_1, \ldots, gx_k)$. En outre, il y a une opération induite de G sur l'ensemble de toutes les parties $\mathcal{P}(\Omega)$ (voir chap. 1, § 5, exercice 4). Posons $g\varnothing = \varnothing$ et $gT = \{gt \mid t \in T\}$ si T est un sous-ensemble non vide de Ω. Les propriétés (i), (ii) se vérifient immédiatement. On comprend facilement que T et gT ont même puissance, de sorte que G induit une opération sur des sous-ensembles de même puissance.

2. Orbites et stabilisateurs des points.— On dit que deux points $x, x' \in \Omega$ sont équivalents par rapport au groupe G opérant sur Ω si $x' = gx$ pour un certain élément $g \in G$. Les propriétés de réflexivité, de symétrie et de transitivité qu'on obtient facilement à l'aide de (i), (ii) (voir n° 1) montrent que nous avons affaire à une relation d'équivalence qui répartit Ω en classes d'équivalence disjointes. On convient de donner à ces classes le nom de *G-orbites.* Il est naturel de

désigner une orbite contenant l'élément $x_0 \in \Omega$ par le symbole $G(x_0)$; ainsi, $G(x_0) = \{gx_0 \mid g \in G\}$. Pourtant, on utilise aussi d'autres désignations qui soulignent les particularités d'une telle ou telle opération de G sur Ω. La notion d'orbite provient de la géométrie. Si, par exemple, $G = SO(2)$ est le groupe de rotations sur le plan autour de l'origine O, l'orbite d'un point P sera la circonférence de centre O, passant par P, alors que l'ensemble $\Omega = \mathbb{R}^2$ sera la réunion des circonférences concentriques, y compris la circonférence de rayon nul (le point O). Pour nous, la notion d'orbite n'est pas nouvelle non plus. Nous l'avons utilisée au chapitre 4 lors de la décomposition de la permutation $\pi \in S_n$ en un produit de cycles indépendants. A titre de G, nous avions pris le groupe cyclique $\langle\pi\rangle$.

Soit x_0 un point fixe de Ω. Considérons l'ensemble

$$\mathrm{St}(x_0) = \{g \in G \mid gx_0 = x_0\} \subset G.$$

Puisque $ex_0 = x_0$ et $g, h \in \mathrm{St}(x_0) \Rightarrow gh^{-1} \in \mathrm{St}(x_0)$, alors $\mathrm{St}(x_0)$ est un sous-groupe de G. On l'appelle *stabilisateur* (ou *sous-groupe stationnaire*) dans G du point $x_0 \in \Omega$ et on le désigne souvent par le symbole G_{x_0}. Pour le groupe $SO(2)$ opérant sur l'espace $\mathbb{R}^2$, considéré plus haut, nous avons $\mathrm{St}(O) = SO(2)$ et $\mathrm{St}(P) = e$, si $P \neq O$. Dans le cas général

$$gx_0 = g'x_0 \Leftrightarrow g^{-1}g' \in \mathrm{St}(x_0) \Leftrightarrow g' \in g\,\mathrm{St}(x_0).$$

Ainsi donc, les classes à gauche $g\,\mathrm{St}(x_0)$ du groupe G suivant le stabilisateur $\mathrm{St}(x_0)$ sont en correspondance biunivoque avec les points de l'orbite $G(x_0)$. En particulier,

$$\mathrm{Card}\, G(x_0) = \mathrm{Card}\,(G/\mathrm{St}(x_0)) = (G : \mathrm{St}(x_0)). \tag{1}$$

Ici, comme précédemment, $G/\mathrm{St}(x_0)$ est l'ensemble quotient de G par le sous-groupe $\mathrm{St}(x_0)$ et $(G: \mathrm{St}(x_0))$ est l'indice du sous-groupe $\mathrm{St}(x_0)$ de G. La puissance $\mathrm{Card}\, G(x_0)$ est souvent appelée *longueur de la G-orbite* du point x_0.

De (1) et du théorème de Lagrange il résulte que la *longueur de toute orbite par rapport à un groupe fini G est un diviseur de l'ordre du groupe.* ■

Fixons aussi notre attention sur le fait que le point x_0, figurant au second membre de la relation (1), peut être remplacé par tout point $x_0' \in G(x_0)$. En effet,

$$\mathrm{Card}\, G(x_0) = \mathrm{Card}\, G(x_0') = (G : \mathrm{St}(x_0')).$$

Une assertion plus forte, relative aux stabilisateurs, concerne la propriété suivante. Soit $x_0' = gx_0$. Alors

$$\mathrm{St}(x_0')\, gx_0 = \mathrm{St}(x_0')\, x_0' = x_0' = gx_0,$$

d'où

$$g^{-1}\mathrm{St}(x_0')\, gx_0 = x_0, \qquad \text{c'est-à-dire} \qquad g^{-1}\mathrm{St}(x_0')\, g \subset \mathrm{St}(x_0).$$

De même

$$g \operatorname{St}(x_0)\, g^{-1} \subset \operatorname{St}(x_0'),$$

puisque

$$\operatorname{St}(x_0)\, g^{-1}x_0' = \operatorname{St}(x_0)\, x_0 = x_0 = g^{-1}x_0'.$$

On a donc l'égalité

$$\operatorname{St}(x_0') = g \operatorname{St}(x_0)\, g^{-1} = \{ghg^{-1} \mid h \in \operatorname{St}(x_0)\}.$$

Dans l'esprit de l'exemple 1 considéré ci-dessous, deux sous-groupes $H, H' \subset G$ sont dits *conjugués* si $H' = gHg^{-1}$ pour un certain $g \in G$. Enonçons les résultats obtenus sous la forme d'un théorème.

Théorème 1.— *Soit G un groupe opérant sur un ensemble Ω. Si deux points x_0, $x_0' \in \Omega$ sont dans une même orbite, leurs stabilisateurs sont conjugués:*

$$x_0' = gx_0 \Rightarrow \operatorname{St}(x_0') = g \operatorname{St}(x_0)\, g^{-1}.$$

Si G est un groupe fini et

$$\Omega = \Omega_1 \cup \Omega_2 \cup \ldots \cup \Omega_r$$

est une partition de Ω en un nombre fini d'orbites de représentants $x_1, x_2, \ldots, x_r$, alors

$$|\Omega| = \sum_{i=1}^{r} (G : \operatorname{St}(x_i)). \quad \blacksquare \qquad (2)$$

La formule (2) est à la base de nombreuses applications de la « méthode des orbites » aux groupes finis.

3. Exemples d'opérations des groupes sur les ensembles.— Nous ne nous étendrons que sur des exemples ayant trait à la théorie des groupes proprement dite.

Exemple 1 (*opération par conjugaison*). — L'opération de tout élément $g \in G$ sur $\Omega = G$ est définie par la formule

$$x \mapsto I_g(x) = gxg^{-1}, \quad \forall\, x \in G.$$

On pourrait écrire $g \circ x = gxg^{-1}$, mais nous préférons la désignation que nous avons déjà utilisée (voir chap. 4, § 3, nº 2) pour l'automorphisme intérieur I_g correspondant à l'élément $g \in G$.

L'opération de g, identifiée avec l'opération $I_g \in \operatorname{Inn}(G)$, s'appelle *conjugaison*. Le noyau de l'application $I : g \mapsto I_g$ est le centre du groupe G:

$$Z(G) = \{z \in G \mid I_g(z) = z, \quad \forall\, g \in G\} = \{z \in G \mid zg = gz, \quad \forall\, g \in G\}.$$

L'orbite de l'élément $x \in G = \Omega$, désignée ici par le symbole x^G, s'appelle *classe des éléments conjugués*, ou tout simplement *classe*

de conjugaison contenant x. Si $a, b \in x^G$, on écrit parfois $a \underset{G}{\sim} b$. Pour le stabilisateur St (x), appelé dans ce cas *centralisateur* de l'élément x, on utilise le plus souvent la désignation $C(x)$ (ou $C_G(x)$ s'il est nécessaire de porter l'accent sur le groupe G).

D'après la remarque faite à la fin du n° 1, l'opération de conjugaison agit aussi sur les sous-ensembles et les sous-groupes de G. On dit que deux sous-ensembles $H, T \subset G$ sont *conjugués*, si $T = gHg^{-1}$ pour un certain $g \in G$. Soit H un sous-groupe de G. On convient de dire que

$$N(H) = \mathrm{St}(H) = \{g \in G \mid gHg^{-1} = H\}$$

est le *normalisateur* du sous-groupe H dans G. En particulier, $H \lhd G$ (H est un sous-groupe distingué de G) si $N(H) = G$, ce qui est en accord avec les définitions données au chapitre 4. Conformément à la relation (1), la *longueur de l'orbite* H^G *(le nombre de sous-groupes conjugués de* H*) coïncide avec l'indice du normalisateur* $N(H)$ *dans* G. ■

Soient G un groupe fini et $x_1^G, \ldots, x_r^G$ ses classes de conjugaison dont les q premières sont à un élément:

$$x_i^G = \{x_i\}, \quad i = 1, \ldots, q \ (x_1 = e).$$

Alors, $Z(G) = \{x_1, x_2, \ldots, x_q\}$ et les relations (1) et (2) s'écrivent sous la forme

$$|x_i^G| = (G : C(x_i)); \tag{1'}$$

$$|G| = |Z(G)| + \sum_{i=q+1}^{r} (G : C(x_i)). \tag{2'}$$

Supposons par exemple $G = S_3$. Alors $r = 3$, $q = 1$ (c'est-à-dire $Z(S_3) = e$) et

$$S_3 = \{e\} \cup \{(12), (13), (23)\} \cup \{(123), (132)\}$$

est la partition de S_3 en classes de conjugaison. Les cardinaux de ces classes (les longueurs des orbites) divisent $6 = |S_3|$, comme le prescrit la relation (1'). La relation (2') conduit immédiatement à l'assertion intéressante suivante:

THÉORÈME 2.— *Le centre* $Z(G)$ *de tout* p*-groupe fini* G *(groupe d'ordre* $p^n > 1$*, où* p *est un nombre premier) diffère du sous-groupe unité.*

DÉMONSTRATION. — Si G est un groupe abélien, on a $G = Z(G)$ et il n'y a rien à démontrer. Dans le cas contraire, $r > q$, $(G : C(x_i)) = p^{n_i}$, $n_i \geqslant 1$ pour $i > q$, et la relation (2') mise sous la forme

$$p^n = |Z(G)| + \sum_{i=q+1}^{r} p^{n_i},$$

montre que $|Z(G)|$ est divisible par p. ■

L'existence d'un p-groupe non abélien s'établit sans peine. Il suffit de considérer le groupe des matrices triangulaires supérieures

$$P = \left\{ \begin{Vmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{Vmatrix} \mid a, b, c \in Z_p \right\}$$

à coefficients dans un corps commutatif fini à p éléments.

Exemple 2 (*translation*).— L'application $L_a : G \to G$ qui est définie par la formule $L_a(g) = ag$ et que nous avons utilisée lors de la démonstration du théorème de Cayley (voir chap. 4, § 3) est généralement appelée *translation à gauche* définie par a. Puisque $eg = g$ et $(ab)g = a(bg)$, les translations à gauche déterminent une opération de G sur lui-même qui induit une opération sur l'ensemble des parties du groupe G. Soient, en particulier, H un sous-groupe et G/H l'ensemble des classes à gauche gH, $g \in G$.

Il est clair que l'application

$$(x, gH) \mapsto x(gH) = (xg)H$$

définit une opération L^H du groupe G sur G/H. Le noyau Ker L^H de cette opération est l'ensemble

$$\{x \in G \mid L_x^H(gH) = gH, \ \forall g \in G\} = \{x \in G, \ xgH = gH, \ \forall g \in G\}.$$

En d'autres termes, $x \in \operatorname{Ker} L^H \Leftrightarrow g^{-1}xg \in H$ pour tout $g \in G$ où, ce qui revient au même, $x \in gHg^{-1}$, $\forall g \in G$.

Ainsi

$$\operatorname{Ker} L^H = \bigcap_{g \in G} gHg^{-1}$$

est le plus grand sous-groupe distingué du groupe G, contenu dans H. L'efficacité de l'opération de G sur G/H est équivalente à l'absence de sous-groupe $K \subset H$, $K \neq e$, distingué dans G.

En tout cas, tout sous-groupe H d'indice n de G peut être utilisé pour la représentation $(L^H, G/H)$ du groupe G par les permutations L_x^H sur les classes de G suivant H. Cette représentation (peut-être non exacte, c'est-à-dire n'étant pas un monomorphisme) est beaucoup plus économique que celle obtenue à l'aide du théorème de Cayley.

Exemple 3 (*groupes transitifs*).— Un groupe de permutations $G \subset S_n$ opérant sur l'ensemble $\Omega = \{1, 2, \ldots, n\}$ est dit *transitif* si l'orbite G_i d'un certain point $i \in \Omega$ (et donc de tout point) coïncide avec Ω. En d'autres termes, l'opération $G \times \Omega \to \Omega$ est *transitive* sur Ω si, pour tout couple de points $i, j \in \Omega$, il existe au moins un élément $g \in G$ avec $g(i) = j$.

Soit $\Omega^{[k]}$ l'ensemble des parties ordonnées à k éléments de l'ensemble Ω. Le groupe G opérant sur Ω induit une opération sur $\Omega^{[k]}$; si, dans ces conditions, il y a transitivité sur $\Omega^{[k]}$, G s'appelle groupe *k-transitif* sur Ω. Par exemple, le groupe symétrique S_n est n-transitif sur Ω, et le groupe alterné A_n est $(n - 2)$-transitif.

Tout groupe G opère transitivement sur l'ensemble G/H des classes à gauche de G suivant H (voir exemple 2). En effet, si g_iH, g_jH sont deux classes, alors $g_jg_i^{-1}(g_iH) = g_jH$. C'est d'autant plus étonnant qu'on sache si peu de choses sur les groupes k-transitifs pour $k > 5$. Il existe même une hypothèse (non démontrée), avancée il y a plus de cent ans par C. Jordan, qui affirme que ces groupes ne sont qu'au nombre de deux: S_n et A_n.

Nous nous proposons d'obtenir des résultats quantitatifs intéressants sur les groupes transitifs, dont nous aurons besoin par la suite. Soit G un groupe transitif sur Ω. Le stabilisateur St (i) du point $i \in \Omega$ sera désigné par le symbole G_i. Nous savons (voir théorème 1) que $G_i = g_iG_1g_i^{-1}$, $i = 1, 2, \ldots, n$ $(g_1 = e)$ si $i = g_i(1)$. En outre, les éléments g_i peuvent être choisis comme représentants des classes à gauche de G suivant G_1:

$$G = G_1 \cup g_2G_1 \cup \ldots \cup g_nG_1. \tag{3}$$

En particulier, $|G| = n|G_1|$, ce qui s'accorde avec les résultats généraux relatifs aux longueurs des orbites (voir n° 2).

THÉORÈME 3.— *Soient G un groupe transitif sur Ω et $N(g)$ le nombre de points de Ω qui restent fixes lors de l'opération de tout $g \in G$. Alors*:

(i) $\sum_{g\in G} N(g) = |G|$ *(en divisant les deux membres de l'égalité* (i) *par $|G|$, on obtient qu'« en moyenne » chaque élément laisse fixe un point)*;

(ii) *si G est un groupe 2-transitif, alors*

$$\sum_{g\in G} N(g)^2 = 2|G|.$$

DÉMONSTRATION. (i) — On a

$$\sum_{g\in G} N(g) = \sum_{j=1}^{n} \Gamma(j),$$

où $\Gamma(j)$ est le nombre d'éléments de G laissant fixe le symbole j. Autrement dit, $\Gamma(j) = |G_j|$. Or, en raison de la transitivité, $|G_j| = |g_jG_1g_j^{-1}| = |G_1|$, où g_j sont pris de la décomposition (3). Donc

$$\sum_{g\in G} N(g) = \sum_{j=1}^{n} |G_j| = \sum_{j=1}^{n} |G_1| = n|G_1| = |G|.$$

(ii) La condition que G est 2-transitif signifie que le stabilisateur G_1 opère transitivement sur l'ensemble $\Omega_1 = \Omega \setminus \{1\}$, c'est-à-dire que les G_1-orbites seront $\{1\}$ et Ω_1. Soit $N'(x)$ le nombre de points de Ω_1 qui restent fixes lors de l'opération de $x \in G_1$. La relation (i) appliquée au couple (G_1, Ω_1) donne

$$\sum_{x \in G_1} N'(x) = |G_1|.$$

Puisque $N(x) = 1 + N'(x)$ pour tout $x \in G_1$ (on ajoute le point 1), il vient

$$\sum_{x \in G_1} N(x) = 2|G_1|.$$

On a exactement les mêmes relations pour tous les autres G_j:

$$\sum_{x \in G_j} N(x) = 2|G_j| = 2|G_1|.$$

En sommant sur j, on obtient

$$\sum_{j=1}^{n} \sum_{x \in G_j} N(x) = 2n|G_1| = 2|G|.$$

Ici, $N(x)$ est compté une fois pour chaque sous-groupe G_j qui contient x. Or, x laisse fixes $N(x)$ points, et par conséquent il est contenu exactement dans $N(x)$ sous-groupes G_j. Cela signifie qu'à chaque élément x il correspond dans la somme le terme $N(x)^2$. D'autre part, tout élément $y \in G$ non contenu dans la réunion $\bigcup_j G_j$ permute tous les points, de sorte que $N(y) = 0$. On peut donc écrire la relation

$$\sum_{g \in G} N(g)^2 = \sum_{j=1}^{n} \sum_{x \in G_j} N(x) = 2|G|. \quad \blacksquare$$

4. Espaces homogènes.— Le cas qui présente un intérêt particulier pour la géométrie est celui où Ω est un espace topologique (par exemple, la droite $\mathbb{R}$ ou la sphère S^2), G est un groupe dit continu (ou topologique) et l'opération $(g, x) \mapsto gx$ est soumise à une exigence judicieuse:

(iii) $f(g, x) = gx$ est une fonction continue de deux variables g et x.

Un groupe G opérant sur Ω de manière à vérifier les propriétés (i), (ii) du n° 1 et (iii) s'appelle *groupe des déplacements* de l'espace Ω. Ceci étant, on peut considérer des déplacements qui conservent une métrique quelconque sur Ω. On dit que l'espace Ω est *homogène* si G opère sur Ω transitivement au sens de l'exemple 3, c'est-à-dire si tous les points de Ω appartiennent à une seule et même G-orbite.

Des considérations générales présentées aux n[os] 1 et 2 on peut déduire qu'il y a une correspondance biunivoque entre les points de l'espace homogène Ω et les classes de G suivant l'un des stabilisateurs H. Ceci étant, au déplacement $g \in G$ de l'espace Ω correspond l'application $g'H \mapsto gg'H$ sur l'ensemble G/H.

Proposons-nous de considérer d'un nouveau point de vue l'exemple bien connu du groupe SO(3) que nous avons examiné au § 1. Il est commode de se

représenter le groupe SO(3) comme opérant sur une sphère unité S^2 de dimension deux. Il est évident qu'à tout couple de points P, $Q \in S^2$ correspond un déplacement (une rotation) qui transforme P en Q, c'est-à-dire que S^2 est un espace homogène avec le groupe des déplacements SO(3). Le stabilisateur St (P) de tout point $P \in S^2$ laisse fixe tout l'axe passant par P et par le centre O de la sphère. Par conséquent, St $(P) \cong$ SO(2) est un groupe des rotations du plan perpendiculaire à l'axe OP.

Les éléments du groupe SO(2) étant identifiés avec les points de la circonférence S^1 de rayon unité, le groupe SO(3) peut être représenté sous la forme d'un gâteau dont les feuilles sont des cercles unités « numérotés » par les points de la sphère de dimension deux $\mathrm{SO}(3)/S^1 \approx S^2$. Dans ce cas on dit que l'on a affaire à une *décomposition en fibres* (une projection $p\colon \mathrm{SO}(3) \to S^2$) de *base* S^2 et de *fibre* $p^{-1}(P) \approx S'$, $P \in S^2$. Le sens exact de ces notions étant expliqué dans les cours de Géométrie et de Topologie, nous nous contenterons de ce qui vient d'être dit.

EXERCICES

1. Soient Φ et Φ' des homomorphismes d'un groupe G dans $S(\Omega)$ et $S(\Omega')$ respectivement. Les opérations qu'ils définissent sur Ω et sur Ω' sont dites *équivalentes* s'il existe une application bijective $\sigma\colon \Omega \to \Omega'$ qui rend le diagramme

$$\begin{array}{ccc} \Omega & \xrightarrow{\sigma} & \Omega' \\ {\scriptstyle \Phi_g}\downarrow & & \downarrow{\scriptstyle \Phi'_g} \\ \Omega & \xrightarrow{\sigma} & \Omega' \end{array}$$

commutatif pour tout $g \in G$. Ainsi, $\Phi_g^1 = \sigma\Phi_g\sigma^{-1}$. Démontrer que toute opération transitive du groupe G est équivalente à l'opération de G sur les classes à gauche suivant un certain sous-groupe H. (Indication. Prendre pour H le stabilisateur G_1 du point $1 \in \Omega$, utiliser la décomposition (3) et poser $\sigma(i) = g_iG_1$.)

2. En s'appuyant sur le théorème 2 démontrer que tous les groupes d'ordre p^2 (p est un nombre premier) sont abéliens.

3. Montrer que le centre du groupe P indiqué à la fin de l'exemple 1 est de la forme

$$Z(P) = \left\{ \left\| \begin{matrix} 1 & 0 & c \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix} \right\| \mid c \in Z_p \right\}.$$

Déterminer les classes de conjugaison du groupe P.

(Indication. Porter attention au fait que tous les éléments du groupe P sont de la forme

$$g = A^iB^jC^k, \quad \text{où} \quad A = \left\| \begin{matrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix} \right\|, \quad B = \left\| \begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{matrix} \right\|, \quad C = \left\| \begin{matrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix} \right\|;$$

si $g \notin Z(P)$, alors $C_P(g) = \langle g \rangle Z(G)$, $|C_P(g)| = p^2$.)

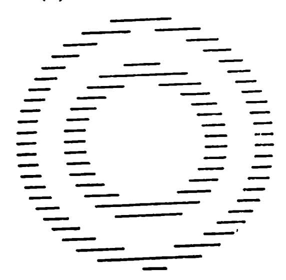
4. Soit n un entier naturel. Mettons-le sous la forme d'une somme $n = n_1 + n_2 + \ldots + n_m$ avec $n_1 \geqslant n_2 \geqslant \ldots \geqslant n_m \geqslant 1$. Le nombre de toutes ces partitions, avec $m = 1, 2, \ldots$, sera désigné par $p(n)$, de sorte que $p(3) = 3$, $p(4) = 5$, etc. La décomposition $\pi = \pi_1\pi_2 \ldots \pi_m$ de chaque permutation $\pi \in S_n$ en un produit de cycles indépendants (voir chap. 4, § 2) définit de façon univoque la partition du nombre n. Montrer que les classes de conjugaison du groupe S_n sont en correspondance bijective avec les partitions du nombre

n. (I n d i c a t i o n. Si $\sigma \in S_n$ et $\pi = \pi_1 \ldots \pi_m$, alors $\sigma\pi\sigma^{-1} = \sigma\pi_1\sigma^{-1} \ldots$ $\ldots \sigma\pi_m\sigma^{-1}$; on a aussi $\sigma \cdot (i_1 i_2 \ldots i_k) \cdot \sigma^{-1} = (\sigma(i_1)\, \sigma(i_2) \ldots \sigma(i_k))$ pour tout cycle $(i_1, i_2 \ldots i_k)$ de longueur k.)

5. Soit $\pi \in S_n$ une permutation s'écrivant sous la forme d'un produit de r cycles de longueur 1, de s cycles de longueur 2, de t cycles de longueur 3, etc., de sorte que $n = r + 2s + 3t + \ldots$ Montrer que la puissance de la classe de conjugaison de S_n, contenant la permutation π, s'exprime par la formule

$$|\pi^{S_n}| = \frac{n!}{1^r r!\, 2^s s!\, 3^t t! \ldots}.$$

6. Soit G un groupe opérant sur un ensemble Ω. On dira qu'un sous-ensemble $\Gamma \subset \Omega$ est *invariant* par rapport à G (ou *G-invariant*) si $gx \in \Gamma$ pour tout $g \in G$ et tout $x \in \Gamma$. Par exemple, les anneaux concentriques sont des ensembles invariants lors de l'opération $SO(2) \times \mathbb{R}^2 \to \mathbb{R}^2$.



Montrer que tout sous-ensemble invariant de Ω est une réunion des orbites et que la G-orbite de tout élément $x \in \Omega$ n'est autre chose que le sous-ensemble invariant le plus petit contenant x.

7. Etant donné un groupe G contenant H comme sous-groupe montrer que l'opération $H \times G \to G$, définie par la translation $(h, g) \mapsto hg$, définit la partition de G en classes à droite de G suivant H.

8. En modifiant la démonstration du théorème 1 obtenir la relation

$$r(G:\Omega) = \frac{1}{|G|} \sum_{g \in G} N(g),$$

où $r(G:\Omega)$ est le nombre d'orbites du groupe des permutations G opérant sur l'ensemble Ω. (I n d i c a t i o n. Dans la somme $\sum N(g)$ chaque élément $x \in \Omega$ est compté $|\mathrm{St}(x)|$ fois. Donc, la contribution que les éléments, situés dans la même orbite que x, apportent à $\sum N(g)$ est égale à $(G: \mathrm{St}(x))\, |\mathrm{St}(x)| = |G|$.)

§ 3. Quelques constructions de la théorie des groupes

Ce paragraphe et surtout son n° 1 présentent certaines difficultés, aussi convient-il d'y revenir plusieurs fois pour pouvoir assimiler, à l'aide des exemples concrets, quelques notions abstraites.

1. Théorèmes généraux sur les homomorphismes des groupes.— Comme nous l'avons vu au chapitre 4, § 4, à chaque sous-groupe distingué K d'un groupe G est associé un certain nouveau groupe G/K appelé groupe quotient du groupe G par K. Ainsi, avec l'épimorphis-

me Φ: $SU(2) \to SO(3)$ décrit au § 1, il est naturel d'introduire le groupe quotient $SU(2)/\{\pm E\}$ et de le comparer à l'image $\operatorname{Im} \Phi = SO(3)$. On se rend compte sans peine que $SU(2)/\{\pm E\} \cong SO(3)$. Or, pour ne pas reprendre chaque fois tous les raisonnements, il est utile d'établir quelques faits généraux relatifs aux sous-groupes, aux homomorphismes et aux groupes quotients. Dans ce qui suit, la notation $K \lhd G$ signifie que K est un sous-groupe distingué de G.

THÉORÈME 1 (théorème fondamental d'homomorphie).— *Soit* $\varphi: G \to H$ *un homomorphisme des groupes de noyau* $K = \operatorname{Ker} \varphi$. *Alors,* K *est un sous-groupe distingué de* G *et* $G/K \cong \operatorname{Im} \varphi$. *Réciproquement, si* $K \lhd G$, *il existe un groupe* H *(à savoir* G/K*) et un épimorphisme* $\pi: G \to H$ *dont le noyau coïncide avec* K *(*π *est souvent appelé application naturelle ou encore homomorphisme naturel).*

DÉMONSTRATION.— Nous savons déjà que $\operatorname{Ker} \varphi = K \lhd G$ (voir chap. 4, § 3, théorème 3). Définissons l'application $\bar{\varphi}: G/K \to H$ en posant

$$\bar{\varphi}(gK) = \varphi(g).$$

Si $g_1K = g_2K$, alors $g_1^{-1}g_2 \in K$, $\varphi(g_1^{-1}g_2) = e$ et donc $\varphi(g_1) = \varphi(g_2)$, ce qui signifie que l'application $\bar{\varphi}$ est définie correctement (c'est-à-dire ne dépend pas du représentant choisi dans la classe). Puisque $\bar{\varphi}(g_1K \cdot g_2K) = \bar{\varphi}(g_1g_2K) = \varphi(g_1g_2) = \varphi(g_1)\varphi(g_2) = \bar{\varphi}(g_1K)\bar{\varphi}(g_2K)$, il vient que $\bar{\varphi}$ est un homomorphisme. En réalité, $\bar{\varphi}$ est un monomorphisme, car $\bar{\varphi}(g_1K) = \bar{\varphi}(g_2K)$ entraîne $\varphi(g_1) = \varphi(g_2)$, d'où $\varphi(g_1^{-1}g_2) = e$, $g_1^{-1}g_2 \in K$ et $g_1K = g_2K$. Il est aussi clair que $\operatorname{Im} \bar{\varphi} = \operatorname{Im} \varphi$. Aussi, $\bar{\varphi}$ est-il l'isomorphisme cherché de G/K sur $\operatorname{Im} \varphi$.

Réciproquement, soit $K \lhd G$. Prenons pour π une fonction qui à tout élément de G associe sa classe à gauche suivant K, c'est-à-dire posons $\pi(g) = gK$. Il est clair que toutes les propriétés requises sont vérifiées. ∎

Il convient de remarquer que la donnée du noyau ne définit pas l'homomorphisme de façon univoque. Par exemple, les automorphismes $g \mapsto g$ et $g \mapsto g^{-1}$ d'un groupe abélien, dont l'ordre est un nombre premier $p > 2$, sont différents bien que leurs noyaux coïncident $(= e)$.

Ayant un homomorphisme $\rho: G \to G_1$ et un sous-groupe $H \subset G$, il est naturel de considérer la restriction $\rho|_H$ et l'image du sous-groupe H par cet homomorphisme. Le théorème qui suit simplifie fortement l'analyse de toutes les situations possibles.

THÉORÈME 2 (premier théorème d'isomorphie).— *Soient* G *un groupe,* H *et* K *ses sous-groupes,* K *étant un groupe distingué de* G. *Alors,* $HK = KH$ *est un sous-groupe de* G *contenant* K. *L'inter-*

section $H \cap K$ est un sous-groupe distingué de H et l'application

$$\varphi: hK \mapsto h\,(H \cap K)$$

est un isomorphisme des groupes

$$HK/K \cong H/H \cap K.$$

Démonstration.— La condition $K \lhd G$ mise sous la forme $gK = Kg$, $g \in G$, signifie en particulier que $hK = Kh$ pour tout $h \in H$. L'ensemble $HK = \{hk \mid h \in H,\ k \in K\}$ est constitué d'un certain nombre de classes hK: $HK = \bigcup_{h \in H} hK$. En y remplaçant hK nous obtenons l'égalité

$$HK = \bigcup_{h \in H} hK = \bigcup_{h \in H} Kh = KH.$$

Il est évident que l'élément unité e, qui est contenu dans H et K, l'est également dans HK. On a aussi $(hk)^{-1} = k^{-1}h^{-1} \in KH = HK$ et donc les inverses de tous les éléments de HK appartiennent, eux aussi, à HK. Enfin, $HK \cdot HK = H \cdot KH \cdot K = H \cdot HK \cdot K = HK$, c'est-à-dire l'ensemble HK est stable pour la multiplication. On voit que le sous-ensemble $HK \subset G$ est un sous-groupe de G.

Puisque $K \subset HK$ et $K \lhd G \Rightarrow K \lhd HK$, il y a lieu de parler du groupe quotient HK/K. Soit $\pi: G \to G/K$ l'épimorphisme naturel et $\pi_0 = \pi\,|_H$ la restriction de π à H. Son image Im π_0 se compose de classes hK, $h \in H$, c'est-à-dire de toutes les classes de G suivant K ayant des représentants dans H. Autrement dit, Im $\pi_0 = HK/K$. Ainsi, nous avons l'épimorphisme

$$\pi_0: H \to HK/K.$$

Son noyau Ker π_0 se compose de $h \in H$ pour lesquels $\pi_0(h) \equiv hK = K$ est l'élément unité de HK/K. Or, $hK = K \Leftrightarrow h \in H \cap K$, d'où Ker $\pi_0 = H \cap K$. Comme tout noyau d'un homomorphisme, $H \cap K$ est un sous-groupe distingué de H (ce qu'on vérifie sans peine d'une façon immédiate).

D'après le théorème fondamental d'homomorphie (théorème 1), la correspondance $\overline{\pi}_0: h\,(H \cap K) \mapsto \pi_0(h) = hK$ établit un isomorphisme $H/H \cap K \cong HK/K$. Puisque $\overline{\pi}_0$ est une application bijective, $\varphi = \overline{\pi}_0^{-1}: hK \mapsto h\,(H \cap K)$ est aussi un isomorphisme des groupes HK/K et $H/H \cap K$. ■

S'il existe un premier théorème d'isomorphie, on doit en avoir un deuxième. Il en est ainsi réellement, mais nous n'allons énoncer que sa variante allégée qui porte un nom spécial.

Théorème 3 (théorème de correspondance).— *Soient G un groupe, H et K ses sous-groupes tels que $K \lhd G$ et $K \subset H$. Alors $\overline{H} = H/K$ est un sous-groupe de $\overline{G} = G/K$ et $\pi^*: H \mapsto \overline{H}$ est une application bijective*

de l'ensemble $\Omega(G, K)$ *des sous-groupes de* G *contenant* K *sur l'ensemble* $\Omega(\overline{G})$ *de tous les sous-groupes du groupe* $\overline{G}$. *Si* $H \in \Omega(G, K)$, *alors* $H \lhd G \Leftrightarrow \overline{H} \lhd \overline{G}$ *et*

$$G/H \cong \overline{G}/\overline{H} = (G/K)/(H/K).$$

DÉMONSTRATION.— Soit $H \in \Omega(G, K)$. De la définition de G/K il est immédiat que H/K est un sous-groupe de G/K. Pour nous assurer que l'application $\pi^*: H \mapsto \overline{H}$ est injective, considérons deux sous-groupes $H_1, H_2 \in \Omega(G, K)$ tels que $H_1/K = H_2/K$. Alors, $h_1 \in H_1 \Rightarrow h_1K = h_2K$, $h_2 \in H_2 \Rightarrow h_1 = h_2K$ et, puisque $K \subset H_2$, on a $h_1 \in H_2$, d'où $H_1 \subset H_2$. On vérifie de même l'inclusion $H_2 \subset H_1$. Par conséquent, $H_1 = H_2$.

Montrons maintenant que l'application π^* est une surjection. Soient $\overline{H} \in \Omega(\overline{G})$ et H l'ensemble des éléments de G dont se composent toutes les classes suivant K qui sont les éléments du groupe $\overline{H} \subset \overline{G}$. Alors, on a en particulier $K \subset H$ et $a, b \in H \Rightarrow aK, bK \in \overline{H} \Rightarrow abK = aKbK \in \overline{H} \Rightarrow ab \in H$ et $a \in H \Rightarrow aK \in \overline{H} \Rightarrow a^{-1}K = (aK)^{-1} \in \overline{H} \Rightarrow a^{-1} \in H$. Par conséquent, H est un sous-groupe de G, et $\overline{H} = H/K$ (H est appelé généralement *image réciproque* dans G du sous-groupe $\overline{H} \subset \overline{G}$).

On a une implication assez évidente $H \in \Omega(G, K)$, $H \lhd G \Rightarrow \overline{H} \lhd \overline{G}$ qui découle formellement des égalités $gK \cdot hK \cdot (gK)^{-1} = ghg^{-1}K = h'K \in \overline{H}$ vérifiées par tous les $g \in G$, $h \in H$. Pour les mêmes raisons, $\overline{H} \lhd \overline{G} \Rightarrow ghg^{-1}K = gK \cdot hK \cdot (gK)^{-1} = h'K \Rightarrow ghg^{-1} \in H \Rightarrow H \lhd G$.

Enfin, dans la situation $H \in \Omega(G, K)$, $H \lhd G$, on peut considérer, d'après ce qui a été démontré, deux épimorphismes naturels

$$\pi: G \to G/K; \; \bar{\pi}: \overline{G} \to \overline{G}/\overline{H}$$

($\bar{\pi}(\bar{g}) = \bar{g}\overline{H}$, où $\bar{g} = gK \in \overline{G}$) et leur composée

$$\sigma = \bar{\pi} \circ \pi: G \to \overline{G}/\overline{H},$$

qui est un épimorphisme défini par la loi $\sigma(g) = \bar{\pi}(\bar{g}) = \overline{g\overline{H}}$. On a $\operatorname{Ker} \sigma = \{g \in G \mid \sigma(g) = \overline{H}\} = \{g \in G \mid \bar{g} \in \overline{H}\} = \{g \in G \mid gK = hK$ pour un certain $h \in H\} = H$. Par suite, d'après le théorème fondamental d'homomorphie, l'application $gH \mapsto \overline{g\overline{H}}$ est un isomorphisme entre G/H et $\overline{G}/\overline{H}$. ∎

EXEMPLE 1. — Soit $n = dm$ un entier naturel ayant un diviseur $d > 1$. Il est évident que $n\mathbb{Z} \subset d\mathbb{Z}$ et que l'application $x \mapsto dx + n\mathbb{Z}$ est un épimorphisme des groupes additifs: $\mathbb{Z} \to d\mathbb{Z}/n\mathbb{Z} = \{di + n\mathbb{Z} \mid i = 0, 1, \ldots, m-1\}$ ayant le noyau $m\mathbb{Z}$. En raison du théorème 1, on a l'isomorphisme

$$\mathbb{Z}_m = \mathbb{Z}/m\mathbb{Z} \cong d\mathbb{Z}/n\mathbb{Z}$$

19—0877

(ce qui est d'ailleurs évident), En utilisant le théorème 3, on trouve

$$\mathbb{Z}/d\,\mathbb{Z} \cong (\mathbb{Z}/n\,\mathbb{Z})/(d\,\mathbb{Z}/n\,\mathbb{Z}), \text{ c'est-à-dire } Z_d \cong Z_n/Z_m.$$

En se rappelant le théorème 5 du chapitre 4, § 3, on arrive à la conclusion que *tous les sous-groupes et tous les groupes quotients d'un groupe cyclique sont encore des groupes cycliques.*

Certes, ce résultat peut être obtenu aussi sans avoir recours aux théorèmes sur les homomorphismes.

EXEMPLE 2. Considérons dans le groupe symétrique S_4 les sous-groupes
$V_4 = \{e, (12)(34), (13)(24), (14)(23)\} \lhd S_4$ (voir exercice 4, § 2),
$S_3 = \{e, (12), (13), (23), (123), (132)\}$
(dans ce cas, S_3 est un stabilisateur du point $i = 4$).

Puisque, évidemment, $S_3 \cap V_4 = e$, on a d'après le théorème 2, pour le sous-groupe $H = S_3V_4$:

$$H/V_4 \cong S_3/S_3 \cap V_4 \cong S_3.$$

En particulier, $|H| = |V_4|\ |S_3| = 24$, c'est-à-dire $H = S_4$. Ainsi, S_4 possède un sous-groupe isomorphe à S_3 et un groupe quotient analogue. En appliquant le théorème 3, nous obtenons la liste des éléments de l'ensemble $\Omega\,(S_4, V_4)$ des sous-groupes de S_4 contenant V_4:

$$\Omega\,(S_4, V_4) = \{V_4, \langle(12)\rangle V_4, \langle(13)\rangle V_4, \langle(23)\rangle V_4, A_4 = \langle(123)\rangle V_4, S_4\}.$$

Fixons notre attention sur le fait que pour tout diviseur d du nombre 24, il existe dans S_4 au moins un sous-groupe d'ordre d. En particulier, il y a exactement quatre sous-groupes $\langle(123)\rangle$, $\langle(124)\rangle$, $\langle(134)\rangle$, $\langle(234)\rangle$ d'ordre 3 et trois sous-groupes $\langle(12)\rangle V_4$, $\langle(13)\rangle V_4$, $\langle(23)\rangle V_4$ d'ordre 8 (ce sont des sous-groupes appelés 3-sous-groupes et 2-sous-groupes de Sylow). Les sous-groupes distingués propres (c'est-à-dire $\neq e$ et S_4) ne sont qu'au nombre de deux: V_4 et A_4.

En effet, si $K \lhd S_4$ et $K \cap V_4 \neq e$, alors $K \supset V_4$, car les éléments de V_4 qui diffèrent de l'unité sont tous conjugués par rapport à S_4. En revenant à l'ensemble $\Omega\,(S_4, V_4)$ nous voyons que $K = V_4$ ou A_4. Si $K \cap V_4 = e$, on a évidemment $KV_4 \neq A_4$. Par ailleurs

$$K \lhd S_4, \quad V_4 \lhd S_4 \Rightarrow KV_4 \lhd S_4,$$

et il ne reste qu'à admettre que $KV_4 = S_4$, $K \cong S_3$ si $K \neq e$. Or, dans ce cas K contient une transposition (car $K \cap V_4 = e$), et toutes les transpositions sont conjuguées dans S_4 et engendrent S_4. D'autre part, elles doivent être contenues dans K puisque $K \lhd S_4$. La contradiction ainsi obtenue signifie que le cas $K \cap V_4 = e$ est impossible.

2. Groupes résolubles. — L'expression

$$[x, y] = xyx^{-1}y^{-1},$$

appelée *commutateur* des éléments x, y d'un groupe G, sert de terme correcteur nécessaire pour intervertir x et y:

$$xy = [x, y]\,yx.$$

Si x et y sont commutables, alors $[x, y] = e$. On comprend intuitivement que plus grand est le nombre de commutateurs différant de e dans le groupe G, plus la loi de multiplication définie sur G diffère de la loi commutative. Soit M l'ensemble de tous les commutateurs de G. On appelle *sous-groupe dérivé* (ou encore *commutant*) du groupe G le sous-groupe G' $(= G^{(1)} = [G, G])$, engendré par l'ensemble M (voir chap. 4, § 2, n° 2):

$$G' = \langle [x, y] \mid x, y \in G\rangle.$$

Bien que $[x, y]^{-1} = yxy^{-1}x^{-1} = [y, x]$ soit un commutateur, le produit de deux commutateurs ne doit pas être nécessairement un commutateur, de sorte que G' se compose de tous les produits possibles de la forme

$$[x_1, y_1] [x_2, y_2] \dots [x_k, y_k] \text{ avec } x_i, y_i \in G.$$

Certes, dans chaque cas concret il est désirable d'avoir une description plus précise du sous-groupe dérivé G'.

EXEMPLE. — Soit $G = S_n$. Le commutateur $[\alpha, \beta] = \alpha\beta\alpha^{-1}\beta^{-1}$ de deux permutations quelconques α, $\beta \in S_n$ est évidemment une permutation paire. Donc, $S'_n \subset A_n$. On a ensuite

$$(ij)(ik)(ij)^{-1}(ik)^{-1} = (ij)(ik)(ij)(ik) = (ijk).$$

Vu que les cycles triples (ijk) engendrent tout le groupe alterné A_n (voir chap. 4, § 2, exercice 8), nous concluons que $S'_n = A_n$.

Remarquons que $S'_n \lhd S_n$ et le groupe quotient S_n/S'_n est abélien.

En revenant à la situation générale, considérons un homomorphisme quelconque des groupes $\varphi: G \to \overline{G}$. Puisque

$$\varphi([x, y]) = \varphi(xyx^{-1}y^{-1}) = \varphi(x)\varphi(y)\varphi(x)^{-1}\varphi(y)^{-1} = [\varphi(x), \varphi(y)],$$

on a $\varphi(G') \subset (\overline{G})'$, et $\varphi(G') = (\overline{G})'$ si φ est un épimorphisme. Soient maintenant K un sous-groupe distingué de G et $\varphi = I_a: x \mapsto axa^{-1}$ est un automorphisme intérieur du groupe G, qui induit un endomorphisme quelconque sur K. Du fait de ce qui précède, $I_a(K') \subset K'$ pour tout $a \in G$; en d'autres termes

$$K \lhd G \Rightarrow K' \lhd G. \qquad (1)$$

En particulier, on a $G' \lhd G$. Démontrons maintenant une assertion générale qui met en évidence le sens interne de la notion de sous-groupe dérivé.

THÉORÈME 4. — *Tout sous-groupe $K \subset G$ contenant le sous-groupe dérivé G' du groupe G est un sous-groupe distingué de G. Le groupe quotient G/G' est abélien, et G' est inclus dans tout sous-groupe distingué K tel que G/K est abélien (en particulier, l'ordre maximal du groupe quotient abélien G/K est égal à l'indice $(G: G')$).*

DÉMONSTRATION. — Si $x \in K$, $g \in G$ et $G' \subset K$, alors $gxg^{-1} = (gxg^{-1}x^{-1})x = [g, x]x \in G'K = K$, de sorte que $K \lhd G$. Les conditions $G' \subset K$, $K \lhd G$, qui sont satisfaites en particulier pour $K = G'$ (voir (1)), entraînent que

$$[aK, bK] = aK \cdot bK \cdot a^{-1}K \cdot b^{-1}K = aba^{-1}b^{-1}K = [a, b]K = K,$$

c'est-à-dire que le commutateur de deux éléments quelconques du groupe quotient G/K est égal à l'élément unité ($= K$). Par conséquent, G/K est un groupe abélien. Réciproquement, si $K \lhd G$ et le groupe quotient est abélien, alors

$$[a, b]\, K = [aK, bK] = K$$

quels que soient $a, b \in G$. Par suite, $[a, b] \in K$ et $G' \subset K$, car G' est engendré par les commutateurs. ■

Remarque.— Nous savons maintenant que tout groupe G contient deux sous-groupes distingués importants: le centre $Z(G)$ et le sous-groupe dérivé G'. Le lien entre ces sous-groupes est assez faible, néanmoins la loi générale est la suivante: plus G « se rapproche » d'un groupe abélien, plus $Z(G)$ est grand et plus G' est petit. Il y a un fait plus intéressant:

Le groupe quotient $G/Z(G)$ d'un groupe non abélien G suivant le centre $Z(G)$ ne peut pas être cyclique.

En effet, si $G/Z(G)$ était un groupe cyclique, alors on aurait $G = \bigcup_i a^i Z(G)$ et tout élément de G serait de la forme $g = a^i z$, $z \in Z(G)$. Dans un tel cas on aurait $[g, h] = [a^i z, a^j z'] = a^{i+j-i-j} \times$ $\times [z, z'] = e$, quels que soient les éléments $g, h \in G$, $G' = e$ et G serait un groupe abélien, ce qui contredit l'hypothèse. ■

Dans G' on peut aussi considérer un sous-groupe dérivé $(G')' =$ $= G''$ appelé *sous-groupe dérivé second* (*deuxième commutant*) du groupe G. En réitérant ce procédé nous pouvons définir le sous-groupe dérivé k-ième $G^{(k)} = (G^{(k-1)})'$. D'après (1), $G^{(k)} \lhd G$ et à plus forte raison $G^{(k)} \lhd G^{(k-1)}$. On obtient ainsi une série de sous-groupes distingués

$$G \rhd G^{(1)} \lhd G^{(2)} \lhd \ldots \rhd G^{(k)} \rhd G^{(k+1)} \rhd \ldots \tag{2}$$

avec des groupes quotients abéliens $G^{(k)}/G^{(k+1)}$.

Un groupe G est dit *résoluble* si la série (2) se termine par un sous-groupe unité, c'est-à-dire si $G^{(m)} = e$ pour un certain indice m, le plus petit, qui s'appelle *degré de résolubilité* du groupe G. Il est évident que tout groupe abélien, notamment tout groupe cyclique, est un groupe résoluble de degré 1. En outre, tout groupe résoluble G de degré de résolubilité m contient un sous-groupe distingué abélien $\neq e$, à savoir $G^{(m-1)}$. Comme le montrent les exemples considérés plus haut, $S_4' = A_4$, $A_4' = V_4$, $V_4' = e$. Par suite, le groupe alterné A_4 est un groupe résoluble de degré 2, et le groupe symétrique S_4 est un groupe résoluble de degré 3.

Les groupes résolubles doivent leur nom à la théorie de Galois dont il a été déjà fait mention plus haut (voir chap. 1, § 2, n° 1). La résolubilité du groupe S_4 et de tous ses sous-groupes est à l'ori-

gine de la résolubilité par radicaux des équations algébriques de degré $n \leqslant 4$. Nous renvoyons pour plus de détails aux ouvrages recommandés au début de la partie II.

3. Groupes simples.— Il existe des groupes $\neq e$ qui coïncident avec leur sous-groupe dérivé et donc ne sont pas résolubles. De plus, nous allons établir maintenant l'existence de groupes non abéliens qui ne contiennent pas du tout de sous-groupes distingués non triviaux ($\neq e$ et G). On convient de donner à ces groupes le nom de groupes simples.

LEMME.— *Tout sous-groupe distingué K d'un groupe G est réunion d'un certain ensemble de classes de conjugaison du groupe G.*

En effet, si $x \in K \lhd G$, on a aussi $gxg^{-1} \in K$ pour tout $g \in G$. Par conséquent, K contient avec chaque élément $x \in K$ toute la classe des éléments conjugués x^G, et $K = \bigcup_{i \in I} x_i^G$. ∎

THÉORÈME 5.— *Le groupe alterné A_5 est un groupe simple.*

DÉMONSTRATION.— En effet, en plus de la permutation unité e, le groupe A_5 comporte 15 éléments $(ij)\,(kl)$ d'ordre 2 (trois éléments de cette forme dans le stabilisateur de chacun de points 1, 2, 3, 4, 5), $20 = 2\binom{5}{3}$ éléments (ijk) d'ordre 3 et $24 = 4!$ éléments $(1i_1i_2i_3i_4)$ d'ordre 5. Les éléments d'ordre 2 sont tous conjugués: ils sont évidemment conjugués dans S_5, la conjugaison pouvant être réalisée par des permutations paires, puisque le stabilisateur (relativement à l'opération de conjugaison) de l'élément (12)(34) comporte une permutation impaire (12). Il en est de même des éléments d'ordre 3. Quant aux éléments d'ordre 5, conjugués dans S_4, ils se répartissent dans le groupe A_4 en deux classes dont les représentants sont (12345) et (12354). En effet, $(45)\,(12345)\,(45)^{-1} = (12354)$, et le centralisateur (= stabilisateur) de l'élément (12345) dans A_5 est le

1	15	20	12	12
e	(12) (34)	(123)	(12345)	(12354)

groupe cyclique d'ordre 5 engendré par cet élément. Ainsi nous avons le tableau dont la ligne inférieure indique les représentants des classes de conjugaison, et la ligne supérieure, les puissances de ces classes.

Soit maintenant K un sous-groupe distingué de A_5. D'après le lemme,

$$|K| = \delta_1 \cdot 1 + \delta_2 \cdot 15 + \delta_3 \cdot 20 + \delta_4 \cdot 12 + \delta_5 \cdot 12,$$

où $\delta_1 = 1$ (car $e \in K$) et $\delta_i = 0$ ou 1 pour $i = 2, 3, 4, 5$. On s'assure aisément que la condition imposée à $|K|$ d'être diviseur de l'ordre $|A_5| = 60$ (théorème de Lagrange) ne laisse que deux possibilités:

a) $\delta_2 = \delta_3 = \delta_4 = \delta_5 = 0$; $K = e$,

b) $\delta_2 = \delta_3 = \delta_4 = \delta_5 = 1$; $K = A_5$,

ce qui démontre que A est un groupe simple. ■

On peut montrer maintenant, par récurrence sur n, que *tous les groupes* A_n, $n \geqslant 5$, *sont simples* (résultat de E. Galois). Puisque les sous-groupes des groupes résolubles sont résolubles ($H \subset G \Rightarrow H^{(k)} \subset G^{(k)}$, $k = 1, 2, \ldots$), le théorème 5 entraîne au moins que le groupe symétrique S_n n'est pas résoluble pour $n \geqslant 5$.

THÉORÈME 6.— *Le groupe des rotations* SO(3) *est un groupe simple.*

DÉMONSTRATION.— En vertu du théorème 3, il suffit de s'assurer que tout sous-groupe distingué K du groupe SU(2) contenant le noyau $\{\pm E\}$ de l'épimorphisme Φ: SU(2) → SO(3) (voir § 1, n° 3) et différant de $\{\pm E\}$ coïncide avec SU(2). La relation (5) du § 1 peut être interprétée d'une autre manière, en disant que chaque classe de conjugaison du groupe SU(2) contient une matrice diagonale $d_\varphi = b_{2\varphi} = \text{diag}\,\{e^{i\varphi}, e^{-i\varphi}\}$. Vu que le sous-groupe K est réunion d'un certain ensemble de classes de conjugaison du groupe SU(2) (voir lemme), on peut, sans restreindre la généralité, poser $d_\varphi \in K$ pour un certain $\varphi > 0$ tel que $\sin\varphi \neq 0$.

K doit également contenir tout commutateur de la forme

$$[d_\varphi, g] = d_\varphi(g d_\varphi^{-1} g^{-1}) = \begin{Vmatrix} e^{i\varphi} & 0 \\ 0 & e^{-i\varphi} \end{Vmatrix} \begin{Vmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{Vmatrix} \begin{Vmatrix} e^{-i\varphi} & 0 \\ 0 & e^{i\varphi} \end{Vmatrix} \begin{Vmatrix} \bar{\alpha} & -\beta \\ \bar{\beta} & \alpha \end{Vmatrix} =$$

$$= \begin{Vmatrix} |\alpha|^2 + |\beta|^2 e^{i2\varphi} & * \\ * & |\alpha|^2 + |\beta|^2 e^{-i2\varphi} \end{Vmatrix},$$

où $|\alpha|^2 + |\beta|^2 = 1$ (voir § 1, relation (3)). Pour la trace de la matrice $[d_\varphi, g]$ on obtient l'expression suivante

$$\text{tr}\,[d_\varphi, g] = 2|\alpha|^2 + |\beta|^2 (e^{i2\varphi} + e^{-i2\varphi}) = 2(1 - 2|\beta|^2 \sin^2\varphi).$$

Ici, $|\beta|$ prend toute valeur comprise dans l'intervalle $[0, 1]$ et $\sin\varphi \neq 0$. De nouveau, en raison de la relation (5) du § 1, il existe une matrice unitaire $h \in$ SU(2) telle que $h\,[d_\varphi, g]\,h^{-1} = d_\psi = \text{diag}\,\{e^{i\psi}, e^{-i\psi}\}$ avec $d_\psi \in K$. Puisque $e^{i\psi}$, $e^{-i\psi}$ sont racines de l'équation caractéristique

$$\lambda^2 + (4|\beta|^2 \sin^2\varphi - 2)\lambda + 1 = 0$$

de la matrice $[d_\varphi, g]$, nous obtiendrons pour ψ tous les points du segment $[0, 2\varphi]$ lorsque $|\beta|$ parcourt les valeurs de 0 à 1. Ainsi donc, K contient tout élément d_ψ et la classe de conjugaison définie par le paramètre ψ, $0 \leqslant \psi \leqslant 2\varphi$. Puisque pour tout $\sigma > 0$ il existe

un nombre naturel n satisfaisant à la condition $0 < \psi = \frac{\sigma}{n} \leqslant 2\varphi$, on peut affirmer que K contient un élément donné d'avance $d_\sigma = d_\psi^n$. ■

Les théorèmes 5 et 6 permettent de conclure à eux seuls que la classe de groupes simples contient aussi bien des groupes finis que des groupes infinis, importants pour les applications. Il peut paraître bien étonnant que jusqu'à présent on ne dispose pas d'une description raisonnable de tous les groupes simples finis et il n'est pas clair si on peut l'obtenir.

4. **Produits de groupes.**— Nous allons examiner maintenant une structure qui permet de construire de nouveaux groupes à partir des groupes donnés et que nous avons déjà rencontrée dans divers cas particuliers.

Nous appellerons *produit direct* (*extérieur*) de deux groupes arbitraires A et B l'ensemble $A \times B$ de tous les couples (a, b), où $a \in A$, $b \in B$, muni de l'opération binaire

$$(a_1, b_1)(a_2, b_2) = (a_1a_2, b_1b_2).$$

A la rigueur, il faudrait écrire $(a_1, b_1) * (a_2, b_2) = (a_1 \circ a_2, b_1 \square b_2)$, où $*$, $\circ$, $\square$ sont les opérations binaires définies respectivement sur A, B et $A \times B$, mais pour simplifier les notations nous conviendrons de désigner toutes les opérations par un point (d'ailleurs, en l'omettant lui aussi). En cas de la notation additive des groupes, par exemple abéliens, il est naturel de parler d'une somme directe $A \oplus B$.

L'ensemble $A \times B$ contient les sous-groupes $A \times e$, $e \times B$ isomorphes respectivement à A et à B (encore une convention: les éléments unités de A et B sont désignés par le même symbole e). Il est évident que l'application $\varphi\colon A \times B \to B \times A$, définie par l'égalité $\varphi((a, b)) = (b, a)$, établit un isomorphisme des groupes $A \times B$ et $B \times A$. Dans le cas où l'on dispose de trois groupes A, B, C, on peut parler des produits directs $(A \times B) \times C$ et $A \times (B \times C)$. En posant $\psi(((a, b), c)) = (a (b, c))$, on s'assure aisément que

$$(A \times B) \times C \cong A \times (B \times C).$$

Les propriétés de commutativité et d'associativité du produit direct nous permettent de parler d'un produit direct de n'importe quel nombre fini de groupes $G_1, G_2, \ldots, G_n$ et d'écrire

$$G_1 \times G_2 \times \ldots \times G_n = \prod_{i=1}^{n} G_i$$

sans indiquer de façon explicite, à l'aide de parenthèses, l'ordre dans lequel sont pris deux à deux les produits directs.

THÉORÈME 7.— *Soit G un groupe contenant des sous-groupes distingués A et B. Si $A \cap B = e$ et $AB = G$, alors $G \cong A \times B$.*

DÉMONSTRATION.— De l'égalité $AB = G$, il résulte que tout élément $g \in G$ s'écrit sous la forme $g = ab$, où $a \in A$, $b \in B$. Si, de plus, $g = a_1 b_1$, $a_1 \in A$, $b_1 \in B$, alors $ab = a_1 b_1 \Rightarrow a_1^{-1} a = b_1 b^{-1} \in \in A \cap B = e$. Par conséquent, $a_1 = a$, $b_1 = b$ et nous arrivons à la conclusion que l'expression $g = ab$ est unique. On a aussi $A \lhd \lhd G \Rightarrow k = a\,(ba^{-1}b^{-1}) = aa' \in A$; $B \lhd G \Rightarrow k = (aba^{-1})\,b^{-1} = = b'b^{-1} \in B$, c'est-à-dire le commutateur $k \in A \cap B = e$ est l'élément unité et donc $ab = ba$.

Définissons maintenant l'application $\varphi: G \to A \times B$ en posant $\varphi\,(g) = (a, b)$ pour tout $g = ab$. Du fait de ce qui précède $\varphi\,(gg') = = \varphi\,(aba'b') = \varphi\,(aa'bb') = (aa', bb') = (a, b)\,(a', b') = \varphi\,(ab) \times \times \varphi\,(a'b') = \varphi\,(g)\,\varphi\,(g')$. En outre $\varphi\,(ab) = (e, e) \Leftrightarrow a = e, b = e$, c'est-à-dire Ker $\varphi = e$. L'épimorphie de φ est évidente. Ainsi donc, φ satisfait à toutes les propriétés d'une application isomorphe des groupes. ■

Un groupe G qui satisfait aux conditions du théorème 7 est appelé *produit direct* (*intérieur*) de ses sous-groupes A, B. A la différence du produit direct extérieur, G contient comme facteurs directs les groupes A, B eux-mêmes et non pas leurs copies isomorphes $A \times e$, $e \times B$. Bien entendu, le produit direct extérieur $G = A \times B$ est, lui aussi, un produit direct intérieur des sous-groupes $A \times e$, $e \times B$ et, ayant acquis une certaine habitude, on peut ne pas distinguer entre les deux produits et utiliser tout simplement l'expression « produit direct ».

Une information sur les homomorphismes des produits directs est contenue dans le théorème suivant :

THÉORÈME 8.— *Soit* $G = A \times B$ *et soit* $A_1 \lhd A$, $B_1 \lhd B$. *Alors* $A_1 \times B_1 \lhd G$ *et* $G/(A_1 \times B_1) \cong (A/A_1) \times (B/B_1)$. *En particulier,* $G/A \cong B$.

DÉMONSTRATION.— Soient $\pi: A \to A/A_1$ et $\rho: B \to B/B_1$ des isomorphismes naturels. Définissons l'application $\varphi: G \to (A/A_1) \times \times (B/B_1)$ par l'intermédiaire de la relation $\varphi\,(ab) = (\pi\,(a), \rho\,(b))$. On vérifie directement que φ est un homomorphisme ayant pour noyau Ker $\varphi = A_1 \times B_1$ et pour image $(A/A_1) \times (B/B_1)$. ■

De même que dans la théorie des espaces vectoriels, on peut ici démontrer sans peine que, si G est un groupe contenant des sous-groupes distingués $G_1, \ldots, G_n$, alors $G \cong \prod G_i$ si, et seulement si, $G = \langle G_1, \ldots, G_n \rangle$ et $G_j \cap \langle G_1, \ldots, \hat{G}_j, \ldots, G_n \rangle = e$ pour tout j (l'accent $\wedge$ mis sur G_j indique que la composante G_j est omise). La même affirmation s'exprime par la propriété suivante : *G est un produit direct de ses sous-groupes distingués $G_1, \ldots, G_n$ si tout élément $g \in G$ s'écrit d'une manière et d'une seule sous la forme $g = = g_1 \ldots g_n$, $g_i \in G_i$.* Le produit direct de n groupes H est encore appelé *puissance directe n-ième* et noté $H^n = H \times \ldots \times H$. On

distingue dans H^n un sous-groupe spécial appelé *diagonale* $\Delta = \{(h, h, \ldots, h) \mid h \in H\}$, isomorphe à H.

Si, dans le théorème 7, on supprime la condition $B \lhd G$, on arrive à la notion de *produit semi-direct*: $G = AB$, $A \cap B = e$, $A \lhd G$ (on écrit parfois $G = A \leftthreetimes B$). Dans cette définition il faudrait introduire encore une description de l'*opération* du sous-groupe B par automorphismes sur le sous-groupe distingué A, ce qu'on fait d'ailleurs dans chaque cas concret.

De nombreux groupes que nous connaissons déjà peuvent être représentés sous la forme de produits directs et semi-directs. Par exemple, S_n est un produit semi-direct du sous-groupe distingué A_n et du groupe cyclique $\langle(12)\rangle$ d'ordre 2: $S_n \cong A_n \leftthreetimes Z_2$. En utilisant les notations de l'exemple 2 du nº 1, on peut écrire $A_4 = V_4 \leftthreetimes \langle(123)\rangle \cong (Z_2 \times Z_2) \leftthreetimes Z_3$; $S_4 = V_4 \leftthreetimes S_3 \cong (Z_2 \times Z_2) \leftthreetimes \leftthreetimes (Z_3 \leftthreetimes Z_2)$. Citons encore un exemple: le groupe $A(1, \mathbb{R})$ des transformations affines $\mathbb{R} \to \mathbb{R}$ (voir chap. 4, § 2, exercice 3) est un produit semi-direct du sous-groupe distingué des translations et du sous-groupe $GL(1, \mathbb{R})$ des transformations qui laissent fixe le point $x = 0$.

5. Générateurs. Relations de définition. — Nous avons déjà examiné au chapitre 4, § 2, la question relative aux systèmes de générateurs d'un groupe G. Nous y revenons pour voir sous un jour nouveau certains groupes connus. Des résultats obtenus au cours du chapitre 4 il s'ensuit que pour les groupes cycliques il n'est pas besoin de composer les tables de Cayley encombrantes. La notation conventionnelle

$$C_n = \langle c \mid c^n = e \rangle \tag{3}$$

fournit toute l'information nécessaire sur le groupe cyclique abstrait C_n d'ordre n; on convient que $C_n = \{e, c, c^2, \ldots, c^{n-1}\}$, avec $c^s c^t = c^{s+t}$ pour $s + t < n$, et $c^s c^t = c^{s+t-n}$ pour $s + t \geqslant n$. D'autre part, tout groupe cyclique est l'image homomorphe d'un et d'un seul groupe $(\mathbb{Z}, +)$.

Un groupe, universel dans ce sens pour tous les produits directs possibles $A = \langle a_1 \rangle \times \ldots \times \langle a_r \rangle$ des groupes cycliques d'ordres $n_1, \ldots, n_r$ (n_i est un entier naturel ou le symbole ∞), sera représenté par la puissance directe r-ième $\mathbb{Z}^r = \mathbb{Z} \oplus \ldots \oplus \mathbb{Z}$ (voir nº 4) ayant pour générateurs

$$z_i = (0, \ldots, 1, \ldots, 0),\ i = 1, 2, \ldots, r,$$

la loi de composition étant définie de la façon suivante:

$$\sum s_i z_i + \sum t_i z_i = \sum (s_i + t_i)\, z_i = (s_1 + t_1, \ldots, s_r + t_r).$$

L'application $z_i \mapsto a_i$, $1 \leqslant i \leqslant r$, se prolonge univoquement en l'homomorphisme des groupes $\varphi: (s_1, s_2, \ldots, s_r) \mapsto a_1^{s_1} a_2^{s_2} \ldots a_r^{s_r}$, de noyau $\operatorname{Ker} \varphi = m_1\mathbb{Z} \oplus \ldots \oplus m_r\mathbb{Z}$ (voir théorème 8), où $m_i = n_i$, si $n_i < \infty$, et $m_i = 0$ si $n_i = \infty$.

Par analogie avec (3), on peut écrire

$$A = \langle a_1, \ldots, a_r \mid a_1^{m_1} = e, \ldots, a_r^{m_r} = e \rangle,$$

en supposant implicitement que les générateurs $a_1, \ldots, a_r$ commutent encore. On convient d'appeler $a_1^{m_1} = e, \ldots, a_r^{m_r} = e$ *relations de définition* du groupe abélien A et $\mathbb{Z}^r$ *groupe abélien libre de rang r* (ou groupe abélien *à r générateurs libres* $z_1, \ldots, z_r$). Il est évident que

$$A \cong \mathbb{Z}^r \Leftrightarrow (a_1^{s_1} \ldots a_r^{s_r} = e \Leftrightarrow s_1 = \ldots = s_r = 0).$$

Si maintenant F_d est un groupe arbitraire engendré par d générateurs $f_1, \ldots, f_d$, tout élément f de ce groupe s'écrit (peut-être de plusieurs manières) sous la forme

$$f = f_{i_1}^{s_1} f_{i_2}^{s_2} \ldots f_{i_k}^{s_k}; \quad i_j \in \{1, 2, \ldots, d\}, \ s_j \in \mathbb{Z}, \tag{4}$$

où $i_j \neq i_{j+1}$, $j = 1, 2, \ldots, k - 1$. On y arrive toujours au moyen de substitutions élémentaires $f_i^s f_i^t = f_i^{s+t}$, $f_i^0 = e$ et $f_j e = e f_j = f_j$.

Si les conditions $f = e \Leftrightarrow s_1 = \ldots = s_k = 0$ sont satisfaites pour chaque f écrit sous la forme (4), on dit que F_d *est un groupe libre engendré par d générateurs libres.* Les éléments du groupe F_d sont généralement appelés *mots de l'alphabet* $\{f_1, f_1^{-1}, \ldots, f_d, f_d^{-1}\}$. La *notation irréductible* (4) *du mot* f et sa *longueur* $l(f) = |s_1| + |s_2| + \ldots + |s_k|$ sont définies de façon unique ; dans le cas contraire, le mot $e = ff^{-1}$ (élément neutre de F_d) aurait une longueur > 0. Pour un d donné, deux groupes libres F_d et G_d engendrés par les générateurs libres respectifs $f_1, \ldots, f_d$ et $g_1, \ldots, g_d$ sont isomorphes : il suffit de poser $\Phi(f_i) = g_i$, $1 \leqslant i \leqslant d$, et de considérer, pour un mot arbitraire f de la forme (4), que

$$\Phi(f) = g_{i_1}^{s_1} g_{i_2}^{s_2} \ldots g_{i_k}^{s_k}$$

(les éléments neutres de F_d et G_d sont désignés par les mêmes symboles). Si, pourtant, G_d n'est pas un groupe libre, Φ ne sera qu'un épimorphisme de noyau $\mathrm{Ker}\, \Phi$ composé de ceux des mots que la substitution $f_i \mapsto g_i$ transforme en élément neutre du groupe G_d. Cette *propriété universelle* (la possibilité de prolonger $f_i \mapsto g_i$, $1 \leqslant i \leqslant d$, en l'épimorphisme $\Phi : F_d \to G_d$ pour tout groupe G_d à d générateurs) peut être prise pour la définition du groupe libre F_d, mais nous ne nous étendrons pas sur ce sujet.

Pour que les groupes libres n'aient pas l'air des êtres mystiques, nous indiquerons encore quelques-unes de leurs réalisations concrètes.

$d = 1$. $F_1 \cong (\mathbb{Z}, +)$ est un groupe abélien libre de rang 1 ou, ce qui revient au même, un groupe cyclique infini.

$d = 2$. Soit $\mathbb{Z}[t]$ un anneau des polynômes à une indéterminée t à coefficients rationnels entiers. Dans le groupe spécial linéaire

SL(2, $\mathbb{Z}$ [t]), considérons le sous-groupe F engendré par les matrices

$$A=\begin{Vmatrix}1 & t\\ 0 & 1\end{Vmatrix},\quad B=\begin{Vmatrix}1 & 0\\ t & 1\end{Vmatrix}.$$

Démontrons que F est un groupe libre. Un simple raisonnement par récurrence sur k montre que l'élément

$$W_k = A^{\alpha_1}B^{\beta_1} \dots A^{\alpha_k} B^{\beta_k},\ \alpha_i,\ \beta_i \neq 0,\ 1 \leqslant i \leqslant k,$$

est de la forme

$$W_k=\begin{Vmatrix}1+\dots+\sigma_k t^{2k} & t(\dots+\sigma_{k-1}\alpha_k t^{2(k-1)})\\ t(\dots+\alpha_1^{-1}\sigma_k t^{2(k-1)}) & 1+\dots+\alpha_1^{-1}\sigma_{k-1}\alpha_k t^{2(k-1)}\end{Vmatrix},$$

où $\sigma_k = \alpha_1\beta_1 \dots \alpha_k\beta_k$ et les points désignent les monômes de degré plus petit par rapport à t. Il est clair que $W_k \neq E$. Un élément arbitraire distinct de l'élément unité du groupe F s'écrit soit sous la forme $B^{\beta}A^{\alpha}$, soit sous la forme $W = B^{\beta}W_kA^{\alpha}$. Si $W = E$, alors $W_k = B^{-\beta}A^{-\alpha}$, ce qui est pourtant impossible (comparer les puissances pour $k > 1$, et effectuer une vérification directe pour $k = 1$).

Un petit raisonnement supplémentaire montre que lors de la substitution $t = m$, où m est un entier arbitraire $\geqslant 2$, le groupe F reste encore libre.

Introduisons maintenant la définition suivante:

Définition.— *Soient F_d un groupe libre à d générateurs libres $f_1, \dots, f_d$; $S = \{w_i,\ i \in I\}$ un sous-ensemble des éléments w_i ($f_1, \dots, f_d$) $\in F_d$ et $K = \langle S^{F_d} \rangle$ le groupe distingué le plus petit de F_d contenant S (l'intersection de tous les sous-groupes distingués contenant S). On dit que le groupe G est défini par d générateurs $a_1, \dots, a_d$ et les relations w_i ($a_1, \dots, a_r$) $= e$, $i \in I$, s'il existe un épimorphisme $\pi : F_d \to G$ de noyau K, tel que π (f_k) $= a_k$, $1 \leqslant k \leqslant d$. Dans ce cas on écrit*

$$G = \langle a_1, \dots, a_d \mid w_i (a_1, \dots, a_d) = e,\ i \in I \rangle$$

et on dit que G est un groupe défini par un ensemble fini de relations si Card $I < \infty$.

Le groupe F_d lui-même est « libre de relations », ce qui explique son appellation. De la définition il résulte que tout groupe H à d générateurs $b_1, \dots, b_d$ qui vérifient les mêmes relations w_i ($b_1, \dots, b_d$) $= e$, $i \in I$, et peut-être certaines autres relations, est une image homomorphe du groupe G. En particulier $|H| \leqslant |G|$.

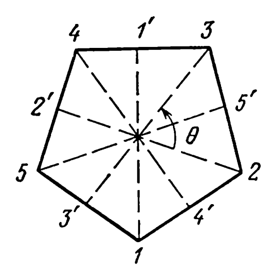
Exemple 1 (*groupe diédral*).— Le groupe $G = \langle a,\ b \mid a^3 = b^2 = abab = e\rangle$ à deux générateurs et à trois relations est d'ordre $|G| \leqslant 6$, car $ba = a^{-1}b^{-1} = (a^3)^{-1}a^2b\cdot(b^2)^{-1} = a^2b$, et les seuls éléments du groupe G sont e, a, a^2, b, ab, a^2b. Puisque les permutations (123), (12), qui engendrent S_3, vérifient les relations $(123)^3 = (12)^2 = (123)\,(12)\,(123)\,(12) = e$, l'application $\varphi: G \to S_3$

définie par la correspondance $a \mapsto (123)$, $b \mapsto (12)$ sera un isomorphisme $G \cong S_3$. Ainsi donc, le groupe symétrique S_3 est défini par deux générateurs et trois relations. Rappelons que S_3 est encore identifié avec le groupe de toutes les transformations de symétrie d'un triangle équilatéral.

Le groupe complet des transformations de symétrie d'un polygone régulier P_n à n côtés s'appelle *groupe diédral* et se note D_n. La rotation

$$\mathcal{A} = \begin{Vmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{Vmatrix}$$

du polygone dans son plan d'un angle $\theta = 2\pi/n$ autour du centre O, situé à l'origine d'un système de coordonnées rectangulaires, engendre un groupe cyclique $\langle\mathcal{A}\rangle$ d'ordre n. D_n contient encore la réflexion $\mathcal{B} = \begin{Vmatrix} 1 & 0 \\ 0 & -1 \end{Vmatrix}$ du polygone P_n par rapport à l'axe passant par le centre et l'un de ses sommets.



Par définition, $\mathcal{B}^2 = e$. Les différentes transformations de symétrie:

$$e,\ \mathcal{A},\ \mathcal{A}^2,\ \ldots,\ \mathcal{A}^{n-1};\ \mathcal{B},\ \mathcal{A}\mathcal{B},\ \ldots,\ \mathcal{A}^{n-1}\mathcal{B}, \tag{5}$$

dont le nombre est égal à $2n$, sont les seuls éléments du groupe D_n. En effet, toute transformation de symétrie est définie par son action sur les sommets $1, 2, \ldots, n$ du polygone P_n. Si une transformation quelconque fait correspondre k à 1, elle doit conserver le même ordre cyclique des sommets, comme le fait $\mathcal{A}^k$, ou bien l'inverser, comme le fait $\mathcal{A}^{k-1}\mathcal{B}$. C'est pourquoi, D_n ne contient aucun élément sauf ceux de (5). Remarquons que la transformation $\mathcal{B}\mathcal{A}$ coïncide avec $\mathcal{A}^{n-1}\mathcal{B}$, car ces deux transformations inversent l'ordre des sommets et transforment 1 en n. On a donc les relations

$$\mathcal{A}^n = e,\ \mathcal{B}^2 = e,\ \ \mathcal{A}\mathcal{B}\mathcal{A}\mathcal{B} = e,$$

ce qui signifie que D_n est l'image homomorphe du groupe

$$G = \langle a,\ b \mid a^n = b^2 = abab = e\rangle.$$

Or, de même que dans le cas de $n = 3$, on obtient $ba = a^{-1}b = a^{n-1}b$, si bien que tout mot de l'alphabet $\{a, a^{-1}, b, b^{-1}\}$ se réduit soit à a^i, soit à a^ib, $0 \leqslant i \leqslant n-1$. Par suite, $|G| \leqslant 2n$ et, du fait de ce qui précède, on doit avoir l'isomorphisme $G \cong D_n$. Nous avons obtenu par là même la donnée du groupe diédral par les générateurs et les relations de définition. Identifions G avec D_n:

$$D_n = \langle a,\ b \mid a^n = e,\ b^2 = e,\ (ab)^2 = e\rangle.$$

Puisque $\langle a\rangle \lhd D_n$ et $D_n/\langle a\rangle$ est un groupe cyclique, le théorème 4 relatif au sous-groupe dérivé D_n' du groupe D_n entraîne l'inclusion $D_n' \subset \langle a\rangle$. Or, $a^2 = aba^{-1}b^{-1} = [a, b] \in D_n'$, et pour n impair $D_n' = \langle a\rangle$, tandis que pour n pair $D_n/\langle a^2\rangle = \langle\bar{a}, \bar{b}\rangle \cong V_4$ est le produit direct de deux groupes cycliques d'ordre 2, d'où $D_n' = \langle a^2\rangle$. Le centre $Z(D_n)$ du groupe D_n et le nombre r de ses classes de conjugaison varient, eux aussi, suivant la parité de n.

Nous donnons ci-dessous des tables prêtes à l'emploi (qui sont d'ailleurs faciles à vérifier):

$n = 2m$. $D'_n = \langle a^2 \rangle$, $(D_n : D'_n) = 4$, $Z(D_n) = \langle a^m \rangle$, $r = m + 3$

1	1	2	...	2	m	m
e	a^m	a	...	a^{m-1}	b	ab

$n = 2m + 1$. $D'_n = \langle a \rangle$, $(D_n : D'_n) = 2$, $Z(D_n) = e$, $r = m + 2$

1	2	...	2	2	n
e	a	...	a^{m-1}	a^m	b

Les représentants des classes de conjugaison sont indiqués dans la ligne inférieure, et les puissances de ces classes dans la ligne supérieure.

Il importe de souligner que la forme des relations de définition (de leurs premiers membres dans $w_i = e$) dépend fortement du choix du système de générateurs du groupe. Par exemple, le groupe diédral D_n est engendré par n'importe quelles transformations par rapport à deux droites qui se coupent sous un angle π/m. Par conséquent,

$$D_n = \langle g_1, g_2 \mid g = g_2^2 = (g_1 g_2)^n = e \rangle$$

Si l'on part de la donnée précédente, on peut poser $g_1 = ab$, $g_2 = b$.

Exemple 2 (*groupe quaternionien*). — A la différence de l'exemple précédent, nous définirons dès le début le groupe quaternionien Q_8 (cette appellation sera expliquée au chap. 9) par les générateurs et les relations:

$$Q_8 = \langle a, b \mid a^4 = e,\ b^2 = a^2,\ bab^{-1} = a^{-1} \rangle.$$

Comme précédemment, $ba = a^{-1}b = a^3 b$ et, puisque $b^2 = a^2$, tout mot de l'alphabet $\{a, a^{-1}, b, b^{-1}\}$ se réduit à la forme $a^s b^t$, $0 \leqslant s \leqslant 3$, $0 \leqslant t \leqslant 1$, si bien que $|Q_8| \leqslant 8$.

Peut-on affirmer que $|Q_8| = 8$? Oui, mais seulement après avoir présenté un groupe à 8 éléments dont deux générateurs sont liés par les mêmes relations que a, b. Un tel groupe est engendré par les matrises

$$A = \begin{Vmatrix} i & \\ 0 & -i \end{Vmatrix}, \quad B = \begin{Vmatrix} 0 & 1 \\ -1 & 0 \end{Vmatrix} \quad (i = \sqrt{-1}).$$

En effet,

$$A^4 = E,\ B^2 = A^2,\ BAB^{-1} = A^{-1}$$

et

$$\langle A, B \rangle = \left\{ \pm \begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix},\ \pm \begin{Vmatrix} i & 0 \\ 0 & -i \end{Vmatrix},\ \pm \begin{Vmatrix} 0 & 1 \\ -1 & 0 \end{Vmatrix},\ \pm \begin{Vmatrix} 0 & i \\ i & 0 \end{Vmatrix} \right\}.$$

L'application $a \longmapsto A$, $b \longmapsto B$ définit l'isomorphisme $Q_8 \cong \langle A, B \rangle$. Remarquons que $a^2 \in Z(Q_8)$ et, puisque le groupe quotient par le centre d'un groupe non abélien ne peut pas être cyclique (voir remarque au nº 2), on a $\langle a^2 \rangle = Z(Q_8)$. Tous les groupes d'ordre 4 étant abéliens, $Q_8/Z(Q_8) \cong V_4$ est le produit direct de deux groupes cycliques d'ordre 2. Par suite, le sous-groupe

dérivé Q'_8 coïncide avec $Z(Q_8)$ et $(Q_8 : Q'_8) = 4$. Les renseignements sur les classes de conjugaison sont rassemblés dans la table ci-dessous:

1	1	2	2	2
e	a^2	a	b	ab

Les groupes définis par un ensemble fini de relations, dont nous venons de considérer les exemples les plus simples, se rencontrent dans les différentes branches des mathématiques, par exemple comme groupes fondamentaux des variétés. Ce n'est pas étonnant que de nombreux problèmes concernant ces groupes restent encore ouverts.

EXERCICES

1. Rappelons-nous la définition de l'automorphisme intérieur (chap. 4, § 3, nº 2) $I_a : g \mapsto aga^{-1}$ et du groupe $\mathrm{Inn}(G) \subset \mathrm{Aut}(G)$. Montrer que $\mathrm{Inn}(G) \lhd \mathrm{Aut}(G)$ et $\mathrm{Inn}(G) \cong G/Z(G)$, où $Z(G)$ est le centre du groupe G. Le groupe quotient $\mathrm{Out}(G) = \mathrm{Aut}(G)/\mathrm{Inn}(G)$ s'appelle *groupe des automorphismes extérieurs*.

2. Soient H et K deux sous-groupes d'un groupe G. Montrer que $|HK| \cdot |H \cap K| = |H| \cdot |K|$ (analogue de la formule connue en théorie des espaces vectoriels). Montrer ensuite que l'ensemble HK est un sous-groupe si, et seulement si, $HK = KH$; dans le cas où $K \lhd G$ cette condition est automatiquement satisfaite.

3. Montrer que dans un groupe résoluble fini G, il existe une suite de sous-groupes $e = G_0 \subset G_1 \subset \ldots \subset G_n = G$, où $G_{i-1} \lhd G_i$, $1 \leqslant i \leqslant n$, et tout indice $(G_i : G_{i-1}) = p_i$ est un nombre premier.

4. Composons pour le groupe symétrique S_4 la table

1	3	6	8	6
e	(12) (34)	(12)	(123)	(1234)

analogue à celle que nous avons utilisée lors de la démonstration du théorème 5. En s'appuyant sur les mêmes raisonnements, reproduire la description des sous-groupes distingués du groupe S_4 que nous avons donnée dans l'exemple 2 du nº 1.

5. Démontrer que le groupe alterné A_n, $n \geqslant 5$, est simple, en raisonnant suivant le schéma esquissé ci-dessous:

a) dans le sous-groupe distingué $K \lhd A_n$, $K \neq e$, il convient de prendre une permutation $\pi \neq e$ qui laisse invariants le plus grand nombre possible k de symboles de $\Omega = \{1, 2, \ldots, n\}$. Si $k = n - 3$, alors $\pi = (ijk)$ et $K = A_n$ (voir chap. 4, § 2, exercice 8); on peut donc poser $k < n - 3$;

b) si $\pi = (123\ldots)\ldots$ est une décomposition de π en cycles indépendants, la parité de π et la condition $k < n - 3$ entraînent $k \leqslant n - 5$. Il est encore possible que $\pi = (12)(34)\ldots$ se compose de cycles indépendants de longueur 2;

c) dans tous les cas, considérer le commutateur $[\pi, \sigma] = \pi\sigma\pi^{-1}\sigma^{-1} \neq e$, avec $\sigma = (345)$, et vérifier qu'il laisse fixes un nombre de symboles supérieur à k. Cela contredit le choix de k et démontre l'assertion avancée.

6. Montrer que $Z(A \times B) = Z(A) \times Z(B)$.

7. Si $K_1, K_2 \lhd G$, $K_1 \cap K_2 = e$, alors G est isomorphe à un certain sous-groupe de $(G/K_1) \times (G/K_2)$. Est-ce vrai?

8. Soit $K \lhd G = A \times B$. Démontrer que le sous-groupe K est abélien, ou bien l'une des intersections $K \cap A$. $K \cap B$ est non triviale. Donner un exemple de groupe $A \times B$ contenant un sous-groupe distingué non trivial K tel que $K \cap A = e$ et $K \cap B = e$. Par là même, $K \lhd A \times B$ n'entraîne pas, en général, que $K = (K \cap A) \times (K \cap B)$.

9. Le groupe quaternionien Q_8 est-il un produit semi-direct de deux de ses sous-groupes propres?

10. Montrer que $H \lhd Q_8$ pour tout sous-groupe propre $H \subset Q_8$.

11. Montrer que les groupes D_4 et Q_8 ne sont pas isomorphes. (Indication. Calculer le nombre d'éléments d'ordre 2 ou utiliser le résultat de l'exercice 10.)

12. Montrer que Aut $(D_4) \cong D_4$ ($|Z(D_4)| = 2$ implique par suite de l'exercice 1 que $|\text{Out}(G)| = 2$).

13. L'ensemble de toutes les racines complexes p^i, $i = 0, 1, 2, \ldots$, de l'unité forme un groupe infini $C(p^\infty)$. Ce groupe est dit *quasi cyclique*, car tout sous-ensemble fini de ses éléments engendre un groupe cyclique.

Vérifier cette proposition et montrer que

$$C(p^\infty) = \langle a_1, a_2, a_3, \ldots | a_1^p = 1, a_{i+1}^p = a_i, i = 1, 2, 3, \ldots\rangle.$$

1. (J. Monthly 80, nº 9 (1973).) Soit

$$G = \langle a, b \mid aba = ba^2b, a^3 = e, b^{2n-1} = e\rangle,$$

où $n \in \mathbb{N}$. Démontrer que $n = 1$, c'est-à-dire que $b = e$ et que $G = \langle a \mid a^3 = e\rangle$ est en réalité le groupe cyclique d'ordre 3. (Indication. $aba = ba^2b = ba^{-1}b \Longrightarrow ab^2 = aba \cdot a^{-1}b = ba^{-1}b \cdot a^{-1}b = ba^{-1} \cdot aba = b^2a$. On en tire la conclusion que $ab = ba$ et donc, compte tenu des autres relations, $b = e$.)

15. Compléter de détails la définition formelle suivante d'un *groupe libre* F_n à n générateurs. On ajoute le symbole e à l'alphabet $A = \{a_1, a_1^{-1}, \ldots, a_n, a_n^{-1}\}$ comprenant n lettres $a_1, \ldots, a_n$ et leurs « antipodes » $a_1^{-1}, \ldots, a_n^{-1}$. Soit S l'ensemble de tous les « mots » obtenus en écrivant ces $2n + 1$ symboles dans un ordre quelconque en lignes de longueur finie, les symboles pouvant se répéter dans les mots. Par produit uv de deux mots u, v on entend l'adjonction du mot v à la fin du mot u. On appelle inverse de $u = a_{i_1}^{\varepsilon_1} \ldots a_{i_m}^{\varepsilon_m}$, $\varepsilon_k = \pm 1$, $k = 1, \ldots, m$, le mot $u^{-1} = a_{i_m}^{-\varepsilon_m} \ldots a_{i_1}^{-\varepsilon_1}$, $e^{-1} = e$. On définit sur S une relation d'équivalence $\sim$. A savoir, deux mots sont considérés comme équivalents si l'un d'eux est obtenu de l'autre en appliquant un nombre fini de transformations élémentaires suivantes :

$$ee \sim e,$$

$$a_i a_i^{-1} \sim e, \quad a_i^{-1} a_i \sim e,$$

$$a_i e \sim a_i, \quad a_i^{-1} e \sim a_i^{-1},$$

$$e a_i \sim a_i, \quad e a_i^{-1} \sim a_i^{-1}.$$

Chaque classe d'équivalence contient un et un seul mot « irréductible » (le plus court). On définit dans l'ensemble des classes d'équivalence par la relation $\sim$ une opération de multiplication associative (et l'inversion des classes), induite par la multiplication de mots. L'élément unité sera représenté par la classe d'équivalence du mot « vide » e. L'ensemble des classes d'équivalence muni de cette opération de multiplication forme justement le groupe libre F_n à n générateurs $a_1, \ldots, a_n$ (*groupe libre de rang* n).

EXEMPLE. — Un « huit » dont les boucles enveloppent deux poteaux est parcouru dans les divers sens par un petit garçon qui tient un fil et met les spires suivantes au-dessus des spires précédentes. Les trajets parcourus par ce garçon, avec des points de départ et d'arrivée au centre entre les poteaux, peuvent manifestement être interprétés comme éléments du groupe libre F_2 de rang 2.

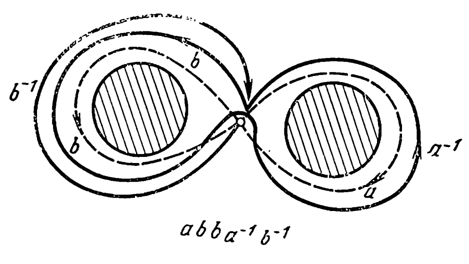


Fig. 19.

Aux mots irréductibles correspondent des fils tendus affranchis des boucles triviales aa^{-1}, $a^{-1}a$, bb^{-1}, $b^{-1}b$. Sur la fig. 19, les tronçons a et a^{-1}, b et b^{-1} ne sont représentés comme géométriquement différents que par souci de clarté. Notre exemple réalise F_2 sous la forme d'une famille de classes de « parcours homotopiquement équivalents » (terminologie topologique) d'une lemniscate. Dans ce sens, le groupe fondamental d'un pétale représenté sur la fig. 21 (page 343) sera le groupe libre F_5.

§ 4. Théorèmes de Sylow

Au chapitre 5, § 3, n° 4, nous avons attiré l'attention sur le fait qu'un groupe fini G d'ordre $|G|$ peut ne pas contenir de sous-groupe d'ordre d divisant $|G|$. Le groupe d'ordre minimal qui vérifie cette propriété est A_4 ($d = 6$).

Puisqu'un groupe simple non abélien ne peut pas contenir de sous-groupes d'indice 2 (du fait qu'ils sont distingués), le groupe alterné A_5 d'ordre 60 ne contient pas en raison du théorème 5 du § 3 de sous-groupes d'ordre 30. Or, en réalité, le groupe A_5 ne contient pas non plus de sous-groupes d'ordres 20 et 15 (Pourquoi? Utilisez les considérations exposées au § 2, n° 3, exemple 2). Sur ce fond, les lois générales établies il y a plus de cent ans par le mathématicien norvégien Sylow sont particulièrement remarquables. Elles concernent les p-groupes (que nous avons rencontrés au § 2) que le groupe G contient comme sous-groupes. L'existence d'un élément d'ordre p dans un groupe abélien, dont l'ordre est divisible par p, a été découverte encore par A. Cauchy.

Soit $|G| = p^n m$, où p est un nombre premier et m est un entier premier avec p. Un sous-groupe $P \subset G$ d'ordre $|P| = p^n$ (s'il existe) sera appelé *p-sous-groupe de Sylow* du groupe G. De même qu'au § 2, n° 3, on entend par $N(P)$ le normalisateur du sous-groupe P de G.

THÉORÈME 1 (premier théorème de Sylow).— *Les p-sous-groupes de Sylow existent.*

THÉORÈME 2 (deuxième théorème de Sylow).— *Soient P et P_1 deux p-sous-groupes de Sylow d'un groupe G. Alors, il existe un élément $a \in G$ tel que $P_1 = aPa^{-1}$. Autrement dit, tous les p-sous-groupes de Sylow sont conjugués.*

THÉORÈME 3 (troisième théorème de Sylow).— *Pour le nombre N_p de p-sous-groupes de Sylow d'un groupe G on a l'égalité $N_p = (G: N(P))$ et la congruence $N_p \equiv 1 \pmod{p}$.*

La démonstration des théorèmes 1 à 3 illustre les méthodes générales et les considérations exposées au § 2. Commençons par le théorème 2.

DÉMONSTRATION du théorème 2.— Ainsi, supposons que les p-sous-groupes de Sylow du groupe G existent et que P soit l'un d'eux. Soit ensuite P_1 un p-sous-groupe arbitraire, non nécessairement de Sylow, du groupe G, qui opère par les translations à gauche sur l'ensemble $G/P = \bigcup_i g_iP$ des classes à gauche de G suivant P (restriction de l'opération de G sur G/P décrite au § 2). Conformément aux résultats du § 2, n° 2, la longueur de toute orbite par rapport à P_1 divise l'ordre $|P_1| = p^k$, $k \leqslant n$. Ainsi,

$$m = \frac{p^n m}{p^n} = \frac{|G|}{|P|} = |G/P| = p^{k_1} + p^{k_2} + \ldots,$$

où p^{k_1}, p^{k_2}, ... sont les longueurs des orbites. Puisque P.G.C.D. $(m, p) = 1$, il existe au moins une orbite ayant la longueur $p^{k_i} = 1$, c'est-à-dire

$$P_1 \cdot aP = aP \tag{1}$$

pour un élément $a = g_i \in G$ (cela ressemble à la démonstration du théorème 2 du § 2). Mettant la relation (1) sous la forme

$$P_1 \cdot aPa^{-1} = aPa^{-1},$$

il vient que

$$P_1 \subset aPa^{-1} \tag{2}$$

(puisque aPa^{-1} est un groupe). En particulier, si P_1 est un p-sous-groupe de Sylow, alors $|P_1| = |P|$, et il résulte par suite de (2) que $P_1 = aPa^{-1}$. ∎

DÉMONSTRATION des théorèmes 1 et 3.— Le théorème 1 peut être interprété comme corollaire du théorème 3, car $N_p \equiv 1 \pmod{p} \Rightarrow N_p \neq 0$, et $N_p \neq 0 \Leftrightarrow S \neq \varnothing$, S étant l'ensemble de tous les p-sous-groupes de Sylow du groupe G.

Quant au théorème 3, l'égalité $N_p = (G: N(P))$ découle directement du fait que les p-sous-groupes de Sylow sont conjugués (théorème 2) et de l'assertion générale sur la longueur de l'orbite H^G

(voir § 2). La congruence $N_p \equiv 1$ (mod. p) sera obtenue quand nous aurons considéré une situation plus générale. A savoir, soit $|G| = p^s t$, où $s \leqslant n$ (t peut être divisible par p), et soit $N_p(s)$ *le nombre de tous les sous-groupes d'ordre* p^s *contenus dans* G. *Il se trouve qu'on a la congruence* $N_p(s) \equiv 1$ (mod. p); en particulier, G contient les sous-groupes de tout ordre p^s, $s = 1, 2, \ldots, n$, et $N_p(n) = N_p$.

Raisonnons comme suit. L'opération du groupe G par les translations à gauche sur lui-même induit, en vertu de la remarque faite à la fin du nº 1, § 2, une opération de G sur l'ensemble

$$\Omega = \{M \subset G \mid |M| = p^s\}$$

de tous les sous-ensembles $\{g_1, \ldots, g_{p^s}\}$ à p^s éléments. Rappelons que $g \cdot \{g_1, \ldots, g_{p^s}\} = \{gg_1, \ldots, gg_{p^s}\}$. L'ensemble Ω se subdivise en G-orbite Ω_i: $\Omega = \bigcup_i \Omega_i$, de sorte que

$$|\Omega| = \sum_i |\Omega_i|, \quad |\Omega_i| = (G : G_i),$$

où $G_i = \{g \in G \mid gM_i = M_i\}$ est le stabilisateur d'un représentant $M_i \in \Omega_i$.

Puisque $G_iM_i = M_i$, alors $M_i = \bigcup_{j=1}^{\nu_i} G_ig_{ij}$ est la réunion de plusieurs classes à droite de G suivant G_i. Par suite $p^s = |M_i| = \nu_i |G_i|$ d'où $|G_i| = p^{s_i} \leqslant p^s$. Dans le cas où $|G_i| < p^s$, on a $|\Omega_i| = p^{s-s_i} t \equiv 0$ (mod. pt); les égalités $|G_i| = p^s$ et $|\Omega_i| = t$ sont équivalentes. On obtient

$$\binom{|G|}{p^s} = |\Omega| \equiv \sum_{|\Omega_i| = t} |\Omega_i| (\text{mod. } pt). \tag{3}$$

Compte tenu de ce qui précède, $|\Omega_i| = t \Rightarrow |G_i| = p^s \Rightarrow M_i = C_ia_i$ ($a_i = g_{i1}$ est un élément de G) et donc $a_i^{-1}M_i = a_i^{-1}G_ia_i = P_i$ est un sous-groupe d'ordre p^s. Les seuls éléments de l'orbite Ω_i sont représentés par un certain nombre de classes à gauche gP_i du groupe G suivant P_i.

Réciproquement, chaque sous-groupe $H \subset G$ d'ordre $|H| = p^s$ conduit à une orbite $\Omega' = \{gH \mid g \in G\}$ de longueur t. Les différents sous-groupes H_i, avec $|H_i| = p^s$, conduisent à des orbites différentes Ω'_i, car $H_i = gH_j$ entraîne $e = gh_j$, d'où $g = h_j^{-1} \in H_j$ et $H_i = H_j$. Ainsi, il existe une correspondance biunivoque entre les sous-groupes d'ordre p^s et les orbites Ω_i de longueur t. La congruence (3) se met sous la forme

$$\binom{|G|}{p^s} \equiv \sum_{|\Omega_i| = t} |\Omega_i| \equiv tN_p(s) \, (\text{mod. } pt), \tag{4}$$

où il faudrait écrire $N_p(s, G)$ pour souligner que $N_p(s)$ est fonction de G.

Jusqu'à présent la nature du groupe G n'a joué aucun rôle. Si l'on prend pour G le groupe cyclique d'ordre $p^s t$, on a pour lui $N_p(s, G) = 1$ (chap. 4, § 3, théorème 5), et donc

$$\binom{|G|}{p^s} \equiv t \cdot 1 \pmod{pt}. \qquad (5)$$

Puisque les premiers membres des congruences (4) et (5) par rapport au même module pt coïncident, on a

$$t \equiv tN_p(s) \pmod{pt},$$

ce qui donne la congruence cherchée $N_p(s) \equiv 1 \pmod{p}$. ■

Bien que nous ayons démontré davantage que cela n'était exigé, nous n'avons point l'intention d'en profiter, en renvoyant aux ouvrages spécialisés tous ceux qui s'y intéressent.

EXEMPLE. — Soit $G = \mathrm{SL}(2, Z_p)$ le groupe de toutes les matrices 2×2, de déterminant 1, sur le corps Z_p à p éléments. De la partition

$$\mathrm{GL}(2, Z_p) = \bigcup_{i=1}^{p-1} \left\| \begin{matrix} i & 0 \\ 0 & 1 \end{matrix} \right\| \mathrm{SL}(2, Z_p)$$

du groupe linéaire complet $\mathrm{GL}(2, Z_p)$ en classes suivant $\mathrm{SL}(2, Z_p)$ il résulte que

$$|\mathrm{GL}(2, Z_p)| = (p-1)\,|\mathrm{SL}(2, Z_p)|. \qquad (6)$$

En considérant $\mathrm{GL}(2, Z_p)$ comme groupe des automorphismes d'un espace vectoriel V de dimension deux sur Z_p, il est facile de trouver l'ordre $|\mathrm{GL}(2, Z_p)|$. En effet, $\mathrm{GL}(2, Z_p)$ opère sur l'ensemble des couples $\{v_1, v_2\}$ de vecteurs de base. Tout vecteur non nul $f_1 \in V$ (ils sont au nombre de $p^2 - 1$) peut servir d'image pour v_1. Quel que soit le choix de f_1, l'image de v_2 peut être représentée par tout vecteur f_2 de $V \setminus \langle f_1 \rangle$ (ces vecteurs sont au nombre de $p^2 - p$). Par suite, $|\mathrm{GL}(2, Z_p)| = (p^2 - 1)(p^2 - p)$, ce qui conduit, conjointement avec (6) à la formule

$$|\mathrm{SL}(2, Z_p)| = p(p^2 - 1).$$

Nous trouvons tout de suite au moins deux p-sous-groupes de Sylow du groupe $\mathrm{SL}(2, Z_p)$:

$$P_1 = \left\{ \left\| \begin{matrix} 1 & \alpha \\ 0 & 1 \end{matrix} \right\| \mid \alpha \in Z_p \right\}, \quad P_2 = \left\{ \left\| \begin{matrix} 1 & 0 \\ \alpha & 1 \end{matrix} \right\| \mid \alpha \in Z_p \right\}.$$

Suivant le théorème 3, on a

$$N_p = (G : N(p)) = 1 + kp > 1.$$

Puisque

$$\left\| \begin{matrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{matrix} \right\| \left\| \begin{matrix} 1 & \alpha \\ 0 & 1 \end{matrix} \right\| \left\| \begin{matrix} \lambda^{-1} & 0 \\ 0 & \lambda \end{matrix} \right\| = \left\| \begin{matrix} 1 & \lambda^2\alpha \\ 0 & 1 \end{matrix} \right\|$$

et, par conséquent, le normalisateur $N(P)$ contient le sous-groupe

$$H = \left\{ \left\| \begin{matrix} \lambda & \alpha \\ 0 & \lambda^{-1} \end{matrix} \right\| \mid \alpha, \ \lambda \in Z_p, \ \lambda \neq 0 \right\}$$

d'ordre $p(p-1)$, il ne reste qu'une seule possibilité

$$N(P) = H, \quad N_p = 1 + p.$$

Entre le groupe

$$\mathrm{SL}(2, Z_2)=\left\{\begin{Vmatrix}1&0\\0&1\end{Vmatrix}, \begin{Vmatrix}1&1\\1&0\end{Vmatrix}, \begin{Vmatrix}0&1\\1&1\end{Vmatrix}, \begin{Vmatrix}0&1\\1&0\end{Vmatrix}, \begin{Vmatrix}1&0\\1&1\end{Vmatrix}, \begin{Vmatrix}1&1\\0&1\end{Vmatrix}\right\}$$

et le groupe symétrique S_3 il s'établit directement l'isomorphisme

$$(1\ 2\ 3)\mapsto\begin{Vmatrix}1&1\\1&0\end{Vmatrix}, \quad (1\ 2)\mapsto\begin{Vmatrix}0&1\\1&0\end{Vmatrix}$$

(la donnée des deux groupes par les générateurs et les relations est la même). Pour $p > 2$, le groupe $G = \mathrm{SL}(2, Z_p)$ a le centre $Z(G) = \{\pm E\}$ d'ordre 2. Le groupe quotient $\mathrm{PSL}(2, Z_p) = G/Z(G)$ qu'il est naturel d'appeler *groupe spécial projectif* (c'est un groupe des transformations de la droite projective $Z_pP^1 = P^1(V) = \{0, 1, \ldots, p-1\} \cup \{\infty\}$) joue depuis Galois un rôle important en algèbre. Cela tient à ce que, pour $p > 3$, le groupe $\mathrm{PSL}(2, Z_p)$ est un groupe simple qui fournit, à côté de A_n, l'un de tout premiers exemples de groupes simples finis.

Revenons maintenant au cas général pour obtenir une précision bien utile des théorèmes de Sylow.

Théorème 4.— *On a les assertions suivantes:*

(i) *un p-sous-groupe de Sylow d'un groupe G est distingué dans G si, et seulement si,* $N_p = 1$;

(ii) *pour qu'un groupe fini G d'ordre* $|G| = p_1^{n_1} \ldots p_k^{n_k}$ *soit le produit direct de ses* p_i*-sous-groupes de Sylow* $P_1, \ldots, P_k$, *il faut et il suffit que tous ces sous-groupes soient distingués dans G.*

Démonstration. (i) — Tous les sous-groupes de Sylow associés à un diviseur premier donné p de l'ordre $|G|$ sont conjugués (deuxième théorème de Sylow), et si P est l'un d'eux, on a $N_p = 1 \Leftrightarrow xPx^{-1} = P, \forall x \in G \Leftrightarrow P \lhd G$.

(ii) Si $G = P_1 \times \ldots \times P_k$ est le produit direct de ses sous-groupes de Sylow, alors $P_i \lhd G$ comme tout autre facteur direct. Par conséquent, la condition d'être distingué est nécessaire.

Soit maintenant $P_i \lhd G$, $1 \leqslant i \leqslant k$, c'est-à-dire $N_{p_i} = 1$. Remarquons tout d'abord que

$$x \in P_i \cap P_j,\ i \neq j \Rightarrow x^{p_i^s} = e,\ x^{p_j^t} = e \Rightarrow x = e.$$

Par suite, $P_i \cap P_j = e$, d'où l'on a

$$[x_i, x_j] = \begin{cases} (x_i x_j x_i^{-1})\, x_j^{-1} = x_j' x_j^{-1} \in P_j \\ x_i\, (x_j x_i^{-1} x_j^{-1}) = x_i x_i' \in P_i \end{cases} \Rightarrow [x_i, x_j] = e,$$

quels que soient $x_i \in P_i$, $x_j \in P_j$, c'est-à-dire les éléments x_i et x_j sont permutables.

Supposons pour un instant que l'élément unité $e \in G$ soit écrit sous la forme $e = y_1 y_2 \ldots y_k$, où $y_i \in P_i$ est un élément d'ordre $a_i = p_i^{b_i}$. En posant $a = \prod\limits_{i \neq j} a_i$ et en utilisant la commutativité

de $y_1, \ldots, y_k$, on obtient

$$e = (y_1 y_2 \ldots y_k)^a = y_1^a y_2^a \ldots y_k^a = y_j^a.$$

Or, a et a_j étant premiers entre eux, $y_j^{a_j} = y_j^a = e \Rightarrow y_j = e$. Cela est vrai pour tout j, donc l'égalité $e = y_1 y_2 \ldots y_k$ ne peut avoir lieu que pour $y_1 = y_2 = \ldots = y_k = e$.

D'autre part, tout élément $x \in G$ d'ordre $r = r_1 r_2 \ldots r_k$, $r_i = p_i^{s_i}$ s'écrit sous la forme

$$x = x_1 x_2 \ldots x_k, \quad |\langle x_i \rangle| = r_i, \quad 1 \leqslant i \leqslant k. \tag{7}$$

Il suffit de poser $x_i = x^{t_i r_i'}$, où les exposants sont déterminés par les conditions

$$r_i' = r/r_i, \qquad 1 = \sum_{i=1}^{k} t_i r_i'.$$

Si maintenant $x = x_1' x_2' \ldots x_k'$ est une autre expression de x sous la forme de produit de p_i-éléments, on aura, du fait de la commutativité des x_i, x_i' ayant des indices inférieurs différents,

$$e = (x_1' x_2' \ldots x_k')(x_1 x_2 \ldots x_k)^{-1} = x_1' x_1^{-1} \cdot x_2' x_2^{-1} \ldots x_k' x_k^{-1},$$

ce qui entraîne, comme il a été montré plus haut, les égalités $x_1' x_1^{-1} = x_2' x_2^{-1} = \ldots = x_k' x_k^{-1} = e$, c'est-à-dire $x_1' = x_1$, $x_2' = x_2, \ldots, x_k' = x_k$.

Ainsi, tout élément du groupe G s'écrit d'une manière et d'une seule sous la forme (7), c'est-à-dire (voir § 3 et § 4, démonstration du théorème 2) $G = P_1 \times \ldots \times P_k$. ■

Remarque. Un p-sous-groupe de Sylow distingué P d'un groupe G est *caractéristique* dans G, c'est-à-dire invariant par tout automorphisme $\varphi \in \mathrm{Aut}(G)$. En effet, $|\varphi(P)| = |P|$ et donc $\varphi(P)$ est un p-sous-groupe de Sylow. Par suite $\varphi(P) = P$ si $N_p = 1$. Il faut aussi noter que les analogues de sous-groupes de Sylow se rencontrent dans des structures algébriques qui ressemblent peu aux groupes finis.

EXERCICES

1. Déterminer le nombre de 5-sous-groupes de Sylow de A_5.
2. Vérifier que l'ensemble P des matrices

$$\pm\begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix}, \quad \pm\begin{Vmatrix} 1 & -1 \\ -1 & -1 \end{Vmatrix}, \quad \pm\begin{Vmatrix} -1 & -1 \\ -1 & 1 \end{Vmatrix}, \quad \pm\begin{Vmatrix} 0 & 1 \\ -1 & 0 \end{Vmatrix}$$

sur Z_3 forme un groupe isomorphe au groupe quaternionien Q_8 et est un 2-sous-groupe de Sylow de $\mathrm{SL}(2, Z_3)$. Montrer que $P \lhd \mathrm{SL}(2, Z_3)$.

3. Montrer que les groupes S_4 et $\mathrm{SL}(2, Z_3)$ ne sont pas isomorphes. Les groupes $\mathrm{PSL}(2, Z_3)$ et A_4 sont-ils isomorphes?

4. Démontrer que tout groupe G d'ordre pq ($p < q$ sont des nombres premiers) est soit cyclique, soit non abélien comprenant un q-sous-groupe de Sylow distingué, ce dernier cas pouvant avoir lieu si, et seulement si, $q - 1$ est divisible par p. En particulier, tous les groupes d'ordre 15 sont cycliques.

5. Obtenir de nouveau (voir chap. 6, § 1) la congruence $(p - 1)! + 1 \equiv 0 \pmod{p}$ pour un p premier par calcul direct du nombre N_p de p-sous-groupes de Sylow contenus dans le groupe symétrique S_p.

§ 5. Groupes abéliens finis

Dans un groupe commutatif (on dit encore groupe abélien) tous les sous-groupes sont distingués. De ce fait évident et du théorème 4 du § 4 il résulte immédiatement que tout groupe abélien A d'ordre $|A| = p_1^{n_1} p_2^{n_2} \ldots p_k^{n_k}$ admet une décomposition

$$A = A(p_1) \times A(p_2) \times \ldots \times A(p_k) \qquad (1)$$

en un produit direct de ses sous-groupes de Sylow $A(p_i)$. Les facteurs directs $A(p_1), \ldots, A(p_k)$ sont souvent appelés *composantes primaires* d'un groupe abélien. La décomposition (1) est définie de façon unique: *toute composante $A(p_i)$ est tout simplement l'ensemble de tous les p_i-éléments* (éléments de A dont les ordres sont les puissances du nombre premier p_i).

Notre but est de représenter le groupe abélien A sous la forme d'un produit direct de groupes les plus simples que sont des groupes cycliques. Si aucune limitation n'est imposée aux ordres des groupes cycliques, la condition d'unicité d'une telle décomposition ne peut pas être satisfaite, comme le montre un exemple simple:

$$A = \langle a \mid a^6 = e \rangle = \langle a^3 \rangle \times \langle a^2 \rangle.$$

Pourtant, l'arbitraire qui est possible dans la décomposition s'avère assez limité, si bien que le résultat final (le théorème 3) est tout à fait satisfaisant.

1. Groupes abéliens primaires.— Nous aurons en vue par la suite que si un groupe abélien A est engendré par ses sous-groupes B, C, on a en réalité $A = BC$; en outre, $A = B \times C$ si, et seulement si, $B \cap C = e$ (voir § 3, nº 4).

A la différence du cas général, un groupe cyclique C_{p^n} d'ordre p^n n'est pas décomposable, c'est-à-dire il ne peut pas être représenté sous la forme d'un produit direct de groupes cycliques d'un plus petit ordre. En effet, si $C_{p^n} = \langle a \rangle$ et $C_{p^i} = \langle a^{p^{n-i}} \rangle$, la suite

$$C_{p^n} \supset C_{p^{n-1}} \supset \ldots \supset C_p \supset e$$

contient tous les sous-groupes du groupe C_{p^n}. Quels que soient les sous-groupes $X \neq e$, $Y \neq e$, choisis parmi ces derniers, leur intersection $X \cap Y \supset C_p$ est non triviale, et donc ils ne peuvent pas être les composantes d'une décomposition directe.

THÉORÈME 1.— *Tout p-groupe abélien fini est un produit direct de groupes cycliques.*

DÉMONSTRATION.— En raisonnant par récurrence et en supposant le théorème démontré pour tous les p-groupes abéliens d'ordre $<p^n$, choisissons dans notre groupe A, $|A| = p^n$, un élément $a \neq e$ d'ordre maximal p^m et passons au groupe quotient $\overline{A} = A/\langle a\rangle$. Puisque $|\overline{A}| = p^{n-m} < p^n$, on a par hypothèse de récurrence

$$\overline{A} = \overline{A}_1 \times \ldots \times \overline{A}_r, \tag{2}$$

où

$$\overline{A}_i = \langle \overline{b}_i\rangle = \langle b_i\langle a\rangle\rangle = \{\langle a\rangle, b_i\langle a\rangle, \ldots, b_i^{p^{m_i}-1}\langle a\rangle\}$$

est un groupe cyclique d'ordre p^{m_i}, $1 \leqslant i \leqslant r$, $m_1 + \ldots + m_r = n - m$. Par définition,

$$\overline{b}_i^{p^{m_i}} = \overline{e} = \langle a\rangle, \text{ c'est-à-dire } b_i^{p^{m_i}} = a^{s_i} \in \langle a\rangle, \tag{3}$$

et bien que chaque élément $x \in A$ soit de la forme

$$x = b_1^{k_1} \ldots b_r^{k_r} \cdot a^k,$$

cette écriture n'est pas en général unique. Nous devons « corriger » les éléments $b_i \in A$ de manière que les exposants s_i dans (3) soient nuls. Il n'est pas difficile de le faire. En nous rappelant que $m_i \leqslant m$ et en élevant les deux membres de la relation (3) à la puissance p^{m-m_i}, nous obtenons

$$e = a^{s_i p^{m-m_i}},$$

d'où $s_i = t_i p^{m_i}$ (voir chap. 4, § 2, théorème 3). Si l'on pose maintenant $a_i = b_i a^{-t_i}$, la relation (3) deviendra

$$a_i^{p^{m_i}} = e, \quad 1 \leqslant i \leqslant r \ (\Leftrightarrow \langle a_i\rangle \cap \langle a\rangle = e), \tag{3'}$$

avec $\overline{a}_i = a_i\langle a\rangle = b_i\langle a\rangle = \overline{b}_i$ et donc $\langle \overline{a}_i\rangle = \overline{A}_i$. De nouveau, on a

$$x = a_1^{k_1} \ldots a_r^{k_r} a^k$$

pour tout $x \in A$, et cette expression est maintenant unique. Dans le cas contraire, on obtient la relation

$$a_1^{\nu_1} \ldots a_r^{\nu_r} a^{\nu} = e, \quad 0 \leqslant \nu_i < p^{m_i}, \quad 0 \leqslant \nu \leqslant p^m$$

(tous les ν_i, ν ne sont pas simultanément nuls) à laquelle l'épimorphisme $A \to \overline{A}$ fait correspondre la relation $\overline{a}_1^{\nu_1} \ldots \overline{a}_r^{\nu_r} = \overline{e}$ qui dans les conditions de la décomposition directe (2) est équivalente au système $\overline{a}_i^{\nu_i} = \overline{e}$, $1 \leqslant i \leqslant r$, ou, ce qui revient au même, à

$a_i^{\nu_i} \in \langle a \rangle$. Or, d'après (3′), cela n'est possible que pour $\nu_i = 0$, et alors $\nu = 0$.

Cela nous conduit à une contradiction qui montre que

$$A = \langle a_1 \rangle \times \ldots \times \langle a_r \rangle \times \langle a \rangle. ■$$

REMARQUE.— La démonstration du théorème 1 que nous venons de donner ressemble à la démonstration géométrique du théorème sur la forme réduite de Jordan de la matrice d'un opérateur linéaire nilpotent (voir Annexe).

Un complément important du théorème 1 est contenu dans le théorème suivant :

THÉORÈME 2.— *Si un p-groupe abélien fini A est décomposé de deux manières en un produit direct de sous-groupes cycliques*

$$A = A_1 \times \ldots \times A_r = B_1 \times \ldots \times B_s,$$

alors $r=s$ et les ordres $|A_i|$ coïncident avec les ordres $|B_j|$ si ces derniers sont ordonnés d'une certaine manière.

DÉMONSTRATION.— Pour $|A| = p$, le théorème est manifestement vrai. Raisonnons par récurrence sur $|A|$. Dès le début, il est commode d'ordonner les composantes A_i et B_j de manière que leurs ordres ne croissent pas :

$$\begin{aligned} &A_i = \langle a_i \rangle, \quad |\langle a_i \rangle| = p^{m_i}, \\ &m_1 \geqslant m_2 \geqslant \ldots \geqslant m_q > m_{q+1} = \ldots = m_r = 1 \,; \end{aligned} \tag{4}$$

$$\begin{aligned} &B_j = \langle b_j \rangle, \quad |\langle b_j \rangle| = p^{n_j}, \\ &n_1 \geqslant n_2 \geqslant \ldots \geqslant n_t > n_{t+1} = \ldots = n_s. \end{aligned} \tag{5}$$

Des relations

$$(xy)^p = x^p y^p, \quad (x^p)^{-1} = (x^{-1})^p,$$

valables pour tout groupe abélien (voir chap. 4, § 1, relation (3)), il résulte que l'ensemble

$$A^p = \{x^p \mid x \in A\}$$

des puissances p-ièmes de tous les éléments de A forme un sous-groupe de A, qui ne dépend d'aucune décomposition de A en un produit direct. D'autre part, si

$$a_1^{i_1} \ldots a_q^{i_q} \ldots a_r^{i_r} = x = b_1^{j_1} \ldots b_t^{j_t} \ldots b_s^{j_s},$$

alors, compte tenu de (4) et (5), on a

$$(a_1^p)^{i_1} \ldots (a_q^p)^{i_q} = x^p = (b_1^p)^{j_1} \ldots (b_t^p)^{j_t}.$$

Donc

$$\langle \tilde{a}_1 \rangle \times \ldots \times \langle \tilde{a}_q \rangle = A^p = \langle \tilde{b}_1 \rangle \times \ldots \times \langle \tilde{b}_t \rangle,$$

où $\tilde{a}_i = a_i^p$, $\tilde{b}_j = b_j^p$ sont des éléments d'ordres respectifs p^{m_i-1} et p^{n_j-1}. Puisque $|A^p| < |A|$, on a, par hypothèse de récurrence, $q = t$ et $m_1 - 1 = n_1 - 1, \ldots, m_q - 1 = n_q - 1$, d'où $m_1 = n_1, \ldots, m_q = n_q$. En remarquant encore que

$$|A_{q+1} \times \ldots \times A_r| = p^{r-q}, \ |B_{t+1} \times \ldots \times B_s| = p^{s-t}, \quad q = t,$$

nous obtenons

$$p^{m_1+\ldots+m_q}p^{r-q} = |A| = p^{m_1+\ldots+m_q}p^{s-q},$$

ce qui signifie que $s = r$, et toutes les assertions du théorème se trouvent démontrées. ∎

Les ordres $p^{m_1}, \ldots, p^{m_r}$ des facteurs cycliques directs sont appelés *invariants* (ou encore *diviseurs élémentaires*) d'un p-groupe abélien fini A. Si deux groupes abéliens A, B possèdent les mêmes invariants, alors

$$A = A_1 \times \ldots \times A_r, \quad B = B_1 \times \ldots \times B_r, \quad A_i \cong C_{p^{m_i}} \cong B_i,$$

et le système d'applications isomorphes $\varphi_i \colon A_i \to B_i$ induit un isomorphisme $\varphi \colon \varphi((a_1, \ldots, a_r)) = (\varphi_1(a_1), \ldots, \varphi_r(a_r))$ entre les groupes A et B. Par suite, le théorème 2 exprime qu'un *groupe A est défini par ses invariants à un isomorphisme près.* En particulier on a le corollaire suivant:

COROLLAIRE.— *Le nombre de groupes abéliens non isomorphes d'ordre p^n est égal au nombre $p(n)$ de partitions*

$$n = n_1 + n_2 + \ldots + n_r, \quad n_1 \geqslant n_2 \geqslant \ldots \geqslant n_r \geqslant 1, \quad 1 \leqslant r \leqslant n. \ ∎$$

Quant à la fonction à valeurs entières $p(n)$, nous l'avons rencontrée lors de la description des classes d'éléments conjugués dans un groupe symétrique S_n (voir § 2, exercice 4). Un groupe abélien d'ordre p^r d'invariants $p, \ldots, p$ est généralement appelé *groupe abélien élémentaire*. Un tel groupe A se caractérise par la condition $A^p = e$. En donnant la préférence à la notation additive, nous remarquons que le groupe abélien A, avec $pA = 0$ (p est un nombre premier), est un espace vectoriel sur un corps commutatif fini $\mathbb{F}_p$ à p éléments. En effet, si l'on identifie les éléments de $\mathbb{F}_p$ avec les classes résiduelles $\bar{k}$ modulo p ($\mathbb{F}_p = Z_p$) et si l'on pose $\bar{k}a = ka$, $a \in A$, on obtient par là même que $\mathbb{F}_p$ opère sur A, ce qui transforme A en un espace vectoriel sur $\mathbb{F}_p$. Cette opération est définie correctement, car $\bar{k} = \bar{k}'$ entraîne $(k - k')\,a = l\,(pa) = 0$. La décomposition de A en une somme directe des sous-groupes cycliques correspond à la décomposition de l'espace vectoriel en une somme di-

recte de sous-espaces de dimension un (théorème sur la base). Ainsi,

$$A \cong Z_p^r = Z_p \oplus \ldots \oplus Z_p.$$

L'exemple qui a été examiné au § 4 donne une idée de l'arbitraire qui règne dans le choix des sous-espaces de base de dimension un même pour $r = 2$: Z_p^2 admet $p(p+1)$ décompositions différentes.

2. Théorème fondamental sur les groupes abéliens finis. — En s'appuyant sur la décomposition (1) et sur son unicité, ainsi que sur les théorèmes 1 et 2, on est conduit directement à l'assertion fondamentale suivante relative aux groupes abéliens:

THÉORÈME 3.— *Tout groupe abélien fini A est le produit direct de sous-groupes cycliques primaires. Deux décompositions quelconques de ce type comportent un même nombre de facteurs de chaque ordre.*

En utilisant la terminologie de la théorie des espaces vectoriels, nous dirons que les éléments $a_1, \ldots, a_r$ d'ordres $d_1, \ldots, d_r$ constituent une *base* du groupe abélien A si chaque élément $x \in A$ s'écrit d'une manière et d'une seule sous la forme

$$x = a_1^{i_1} a_2^{i_2} \ldots a_r^{i_r}, \; 0 \leqslant i_k \leqslant d_k, \quad k = 1, \ldots, r.$$

Bien entendu, dans un tel cas

$$A \langle a_1 \rangle \times \ldots \times \langle a_r \rangle, \quad |A| = d_1 d_2 \ldots d_r, \tag{6}$$

et le théorème 3 est équivalent à l'assertion que dans tout groupe abélien fini A il existe une base dont les éléments sont primaires (c'est-à-dire que leurs ordres d_i sont des puissances de p premiers divisant $|A|$), le système $\{d_1, d_2, \ldots, d_r\}$ ne dépendant pas du choix de la base. Pour cette dernière raison, de même que dans le cas des groupes primaires, les nombres $d_1, \ldots, d_r$ sont appelés *invariants* ou *diviseurs élémentaires* du groupe A. Parfois, on dit encore que $\{d_1, \ldots, d_r\}$ est le *type* de groupe abélien fini A.

Ecrivons tous les invariants, en les disposant en lignes qui correspondent aux différents diviseurs premiers de l'ordre $|A|$:

$$\begin{array}{ll}
p_1^{n_{11}}, \; p_1^{n_{12}}, \; p_1^{n_{13}}, \ldots; & n_{11} \geqslant n_{12} \geqslant n_{13} \geqslant \ldots; \\
p_2^{n_{21}}, \; p_2^{n_{22}}, \; p_2^{n_{23}}, \ldots; & n_{21} \geqslant n_{22} \geqslant n_{23} \geqslant \ldots; \\
\ldots\ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots\ldots \\
p_k^{n_{k1}}, \; p_k^{n_{k2}}, \; p_k^{n_{k3}}, \ldots; & n_{k1} \geqslant n_{k2} \geqslant n_{k3} \geqslant \ldots
\end{array}$$

On peut considérer que toutes les lignes ont même longueur l si certaines d'entre elles sont complétées par les unités.

Les nombres entiers

$$m_j = p_1^{n_{1j}} p_2^{n_{2j}} \ldots p_k^{n_{kj}}, \qquad j = 1, 2, \ldots, l,$$

sont appelés *facteurs invariants* du groupe abélien A. Par la construction même, on a

$$| A | = m_1 m_2 \ldots m_l, \quad m_{j+1} \mid m_j, \qquad j = 1, 2, \ldots, l-1. \tag{7}$$

Passons maintenant de la décomposition (6), mise sous la forme

$$A = (\langle a_{11} \rangle \times \ldots \times \langle a_{k1} \rangle) \times \ldots \times (\langle a_{1l} \rangle \times \ldots \times \langle a_{kl} \rangle),$$

à la décomposition

$$A = \langle u_1 \rangle \times \langle u_2 \rangle \times \ldots \times \langle u_l \rangle \tag{8}$$

à facteurs cycliques directs d'ordres $m_1, m_2, \ldots, m_l$. A cet effet, il suffit de poser

$$u_j = a_{1j} a_{2j} \ldots a_{kj}, \quad 1 \leqslant j \leqslant l,$$

et de se rappeler la proposition énoncée en fin du n° 3 (voir chap. 4, § 2).

Il est évident que, pour un groupe primaire A, les décompositions directes (6) et (8) coïncident, mais dans le cas général la décomposition (8) est plus économique que (6) ($l \leqslant r \leqslant kl$), en outre, elle dégage tout de suite l'élément u_1 d'ordre le plus élevé m_1 ; les ordres de tous les autres éléments du groupe A divisent le premier facteur invariant m_1. Le nombre entier m_1 est encore appelé *exposant* (ou *exponentielle*) du groupe A. *Un groupe abélien A est cyclique si, et seulement si, son exposant coïncide avec l'ordre* $| A |$.

Il reste à ajouter que le problème d'existence d'un groupe abélien A à facteurs invariants donnés $m_1, m_2, \ldots, m_l$ ne se pose pas: il suffit de considérer (dans la notation additive) la somme directe des groupes cycliques $Z_{m_1}, \ldots, Z_{m_l}$.

Enumérons, à titre d'exemple, tous les groupes abéliens d'ordre 16 et 36: $| A | = 16 = 2^4$, $p(4) = 5$: Z_{16}, $Z_8 \oplus Z_2$,

$$Z_4 \oplus Z_4, \; Z_4 \oplus Z_2 \oplus Z_2, \; Z_2^4 = Z_2 \oplus Z_2 \oplus Z_2 \oplus Z_2.$$

$\| A \| = 36 = 2^2 \cdot 3^2$		Diviseurs élémentaires	Facteurs invariants
$Z_4 \oplus Z_9$	$\cong Z_{36}$	4, 9	36
$Z_2 \oplus Z_2 \oplus Z_9$	$\cong Z_{18} \oplus Z_2$	2, 2, 9	18, 2
$Z_4 \oplus Z_3 \oplus Z_3$	$\cong Z_{12} \oplus Z_3$	4, 3, 3	12, 3
$Z_2 \oplus Z_2 \oplus Z_3 \oplus Z_3$	$\cong Z_6 \oplus Z_6$	2, 2, 3, 3	6, 6

Considérons encore un exemple. Ecrivons le groupe $Z_{72} \oplus Z_{84}$ en termes de facteurs invariants. D'abord, exprimons chacun des termes cycliques par les composantes primaires cycliques

$$Z_{72} = Z_8 \oplus Z_9, \; Z_{84} = Z_4 \oplus Z_3 \oplus Z_7.$$

Puis, groupons toutes les composantes primaires

$$Z_{72} \oplus Z_{84} = (Z_8 \oplus Z_4) \oplus (Z_9 \oplus Z_3) \oplus Z_7$$

(somme directe de p-sous-groupes de Sylow). Maintenant il ne reste qu'à dégager un terme cyclique d'ordre maximal dans chaque composante primaire et de répéter ce processus avec les termes restants:

$$Z_{72} \oplus Z_{84} = (Z_8 \oplus Z_9 \oplus Z_7) \oplus (Z_4 \oplus Z_3) = Z_{504} \oplus Z_{12}.$$

Si l'on effectue toutes ces opérations sur le groupe $Z_{36} \oplus Z_{168}$, on obtiendra le résultat analogue.

Par suite

$$Z_{72} \oplus Z_{84} = Z_{36} \oplus Z_{168}$$

(en toute rigueur, au lieu du signe d'égalité on devrait mettre partout le signe $\cong$). Signalons, en particulier, que les exposants des deux groupes sont égaux à 504.

EXERCICES

1. Démontrer le théorème 1 ou même la première partie du théorème 3 sans passer aux groupes quotients (Indication. Le commencement est le même. Puis, au groupe cyclique $\langle a \rangle$ d'ordre maximal m dans A il convient d'ajouter le facteur direct maximal B. Si $\langle a \rangle \times B = A$, tout est démontré. Dans le cas contraire, considérer un élément $c \in A$ non contenu dans $\langle a \rangle \times B$ mais tel que $c^p \in \langle a \rangle \times B$ pour un exposant p premier. Développer les raisonnements ultérieurs dans le groupe $\langle c, \langle a \rangle \times B \rangle$, en cherchant à obtenir pour ce groupe la décomposition $\langle a \rangle \times B'$, $B' \supset B$.)

2. Obtenir la décomposition d'un groupe abélien fini A en composantes primaires sans recourir aux théorèmes de Sylow et sans utiliser, bien entendu, le théorème . En particulier, afin d'obtenir pour $n = d_1 d_2 \ldots d_k$, $d_i = p_i^{e_i}$ (p_i sont des diviseurs premiers différents), la décomposition

$$Z_n \cong Z_{d_1} \oplus Z_{d_2} \oplus \ldots \oplus Z_{d_k}$$

du groupe cyclique $(Z_n, +)$, on peut utiliser l'exemple 1 du nº 1, § 3, ou la proposition du nº 3 du chapitre 4, § 2.

3. Montrer que dans un groupe abélien fini A il existe pour tout $|d| \mid |A|$ au moins un sous-groupe d'ordre d (inversion du théorème de Lagrange).

4. Montrer que les invariants convenablement ordonnés de tout sous-groupe sont diviseurs des invariants d'un groupe abélien.

5. Si $A \oplus A \cong B \oplus B$, où A et B sont des groupes abéliens finis, alors $A \cong B$.

6. Si A, B, C sont des groupes abéliens finis et $A \oplus C \cong B \oplus C$, alors $A \cong B$.

7. Montrer qu'un groupe abélien à facteurs invariants $m_1, \ldots, m_l$ ne peut pas être engendré par un nombre d'éléments inférieur à l.

8. Montrer qu'un groupe abélien fini d'ordre n, qui n'est pas divisible par le carré d'un entier > 1, est cyclique.

9. Enumérer tous les groupes abéliens non isomorphes d'ordre 72.

10. Les groupes $Z_{12} \oplus Z_{72}$ et $Z_{18} \oplus Z_{48}$ sont-ils isomorphes?

CHAPITRE 8

ÉLÉMENTS DE THÉORIE DES REPRÉSENTATIONS

Avant de passer aux définitions rigoureuses de la théorie des représentations linéaires des groupes, nous allons considérer deux problèmes apparentés.

PROBLÈME 1.— Dans un espace $(m + 1)$-dimensionnel V_m des polynômes homogènes à coefficients réels

$$f(x, y) = a_0x^m + a_1x^{m-1}y + \dots + a_{m-1}y^{m-1}x + a_my^m$$

(ou, plus exactement, des fonctions polynomiales $(x, y) \mapsto f(x, y)$) de degré m, on considère l'ensemble des solutions de l'*équation de Laplace* du second ordre

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0 \qquad (*)$$

aux dérivées partielles (voir chap. 6, § 1, exercice 9). L'opérateur de Laplace $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ est linéaire:

$$\Delta = (\alpha f + \beta g) = \alpha \Delta f + \beta \Delta g, \ \forall\ \alpha,\ \beta \in \mathbb{R}.$$

De ce fait les solutions de l'équation (*) forment un sous-espace H_m de l'espace V_m. On vérifie directement que

$$\Delta f = \sum_{k=0}^{m-2} [(m-k)(m-k-1)\, a_k + (k+2)(k+1)\, a_{k+2}]\, x^{m-2-k} y^k.$$

Par suite,

$$\Delta f = 0 \Leftrightarrow (m-k)(m-k-1)\, a_k + (k+2)(k+1)\, a_{k+2} = 0, \quad 0 \leqslant k \leqslant m-2,$$

et tous les coefficients a_k s'expriment par deux d'entre eux, disons par a_0 et a_1. Ainsi, $\dim H_m \leqslant 2$.

Or, on peut indiquer tout de suite deux solutions linéairement indépendantes. En effet, en étendant suivant la linéarité l'action de

l'opérateur Δ sur les polynômes à coefficients complexes, on aura

$$\Delta (x + iy)^m = m (m - 1) (x + iy)^{m-2} + \\ + im\, i (m - 1) (x + iy)^{m-2} = 0, \quad i^2 = -1.$$

En explicitant les parties réelle et imaginaire, on obtient

$$z_m (x, y) \equiv (x + iy)^m = u_m (x, y) + iv_m (x, y),$$

d'où

$$\Delta u_m + i\Delta v_m = \Delta z_m = 0 \Rightarrow \Delta u_m = 0, \quad \Delta v_m = 0.$$

Ainsi,

$$H_m = \langle u_m (x, y), v_m (x, y) \rangle_{\mathbb{R}}.$$

Maintenant, en interprétant x, y comme coordonnées d'un vecteur de l'espace euclidien $\mathbb{R}^2$ muni d'un système fixe de coordonnées rectangulaires, proposons-nous d'envisager ce qui se passera lors de la rotation du plan $\mathbb{R}^2$ autour de l'origine d'un angle quelconque θ (transformation orthogonale de coordonnées):

$$x' = \Phi_\theta (x) = x \cos \theta - y \sin \theta,$$
$$y' = \Phi_\theta (y) = x \sin \theta + y \cos \theta.$$

La règle de dérivation des fonctions composées, connue en Analyse (et facile à vérifier pour les polynômes), donne

$$\frac{\partial^2 f}{\partial x'^2} = \frac{\partial^2 f}{\partial x^2} \cos^2 \theta - 2 \frac{\partial^2 f}{\partial x\, \partial y} \cos \theta \cdot \sin \theta + \frac{\partial^2 f}{\partial y^2} \sin^2 \theta,$$
$$\frac{\partial^2 f}{\partial y'^2} = \frac{\partial^2 f}{\partial x^2} \sin^2 \theta + 2 \frac{\partial^2 f}{\partial x\, \partial y} \cos \theta \cdot \sin \theta + \frac{\partial^2 f}{\partial y^2} \cos^2 \theta,$$

d'où

$$\frac{\partial^2 f}{\partial x'^2} + \frac{\partial^2 f}{\partial y'^2} = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

Cela signifie que l'équation (*) reste invariante pour un changement orthogonal des variables ou, comme on pourrait le dire encore, pour l'opération du groupe $SO(2) = \{\Phi_\theta\}$. En particulier, les polynômes $u_m (x', y')$, $v_m (x', y')$ seront solutions de l'équation (*) et, comme telles, ils s'exprimeront linéairement en fonction de $u_m (x, y)$ et $v_m (x, y)$. Ainsi donc, le groupe $SO(2)$ opère sur l'espace des solutions de l'équation de Laplace. On dit dans ce cas que l'on a affaire à une représentation réelle linéaire de dimension 2

$$\Phi^{(m)} : \Phi_\theta \mapsto \Phi^{(m)} (\theta)$$

du groupe $SO(2)$.

En revenant aux polynômes complexes, nous remarquons que

$$x' + iy' = xe^{i\theta} + iye^{i\theta} = e^{i\theta} (x + iy),$$
$$(x' + iy')^m = e^{im\theta} (x + iy)^m.$$

En gardant pour l'opérateur linéaire complexifié Φ^m (θ) son ancienne désignation, on aura

$$\Phi^{(m)}(\theta): z_m \mapsto z'_m = e^{im\theta} z_m.$$

Les représentations $\Phi^{(m)}: \Phi_\theta \mapsto e^{im\theta}$, $m \in \mathbb{Z}$, dites unitaires de dimension 1 du groupe SO(2), jouent un rôle important en Analyse.

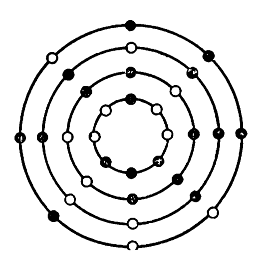


Fig. 20.

Signalons que l'opération Φ induit l'opération du groupe SO(2) sur tout l'espace V_m et, de ce point de vue, H_m est un sous-espace invariant de V_m.

PROBLÈME 2.— L'évaluation du nombre de substances organiques possibles, par exemple en chimie des hydrocarbures cycliques, se ramène à un problème abstrait de tous les jours. Combien de colliers différents de longueur n peut-on confectionner avec un stock illimité de perles de q couleurs différentes ?

Essayons (après G. Polya) de répondre à cette question en considérant que les colliers sont orientés, c'est-à-dire qu'un collier retourné n'est pas en général identifié avec le collier initial. Remarquons que le nombre total de tronçons de fil portant n perles enfilées est égal à q^n (nombre de mots de longueur n contenus dans un demi-groupe libre à q générateurs). Sur l'ensemble Ω_n de ces tronçons opère un groupe cyclique $\langle\sigma\rangle$ d'ordre n, à élément générateur $\sigma = (12 \ldots n) \in S_n$ qui permute circulairement les perles sur chaque tronçon. Il est naturel de considérer comme collier une $\langle\sigma\rangle$-orbite de tronçon ou, si l'on veut, un certain ensemble de cercles concentriques (fig. 20). Cette deuxième interprétation est plus suggestive. Elle est liée à l'isomorphisme

$$\Phi: \sigma \mapsto \Phi(\sigma) = \begin{Vmatrix} \cos\frac{2\pi}{n} & -\sin\frac{2\pi}{n} \\ \sin\frac{2\pi}{n} & \cos\frac{2\pi}{n} \end{Vmatrix}$$

que nous avons déjà rencontré et auquel nous donnerons plus tard le nom de représentation linéaire réelle de dimension 2 du groupe $\langle\sigma\rangle$. Le nombre cherché r de colliers s'exprime par la formule donnée au

chapitre 7, § 2, exercice 8:

$$r=\frac{1}{n}\sum_{k=0}^{n-1} N(\sigma^k).$$

Si $d \mid n$, l'élément σ^d d'ordre n/d laisse invariants ceux des tronçons (et donc les colliers) qui se partagent en d périodes de longueur n/d (voir à ce propos chap. 4, § 2, exercice 13). Par conséquent, $N(\sigma^d) = q^d$ et $N(\sigma^k) = q^{\text{P.G.C.D.}(n,\,k)}$. A la quantité $N(\sigma^k)$, avec P.G.C.D. $(n, k) = d$, correspond dans la somme $\sum N(\sigma^k)$ exactement $\varphi\left(\frac{n}{d}\right)$ termes (φ est une fonction d'Euler). Cela signifie que

$$r=\frac{1}{n}\sum_{d\mid n}\varphi\left(\frac{n}{d}\right)q^d.$$

Le passage à des colliers réellement distincts (non orientés) est lié à des identifications supplémentaires des éléments de Ω_n par l'intermédiaire d'une représentation linéaire ordinaire de dimension deux du groupe diédral D_n. Nous laissons au lecteur le soin de le faire.

Ce n'est pas seulement dans les exemples qui viennent d'être considérés, mais également dans des problèmes physiques réels que les représentations linéaires des groupes apparaissent de façon spontanée comme réflexion d'une symétrie ou d'une autre. Respectivement, les idées et le langage de la théorie des représentations sont bien naturels. C'est ainsi que les exemples traités au § 1 concernent des problèmes bien connus, et il semble à première vue qu'ils ne donnent rien de nouveau. Mais le fait qu'ils apparaissent « sous un même toit » doit, à lui seul, inciter à des raisonnements bien utiles.

La théorie des représentations poursuit deux buts: 1) un but purement mathématique qui est dicté en partie par le désir d'utiliser un appareil supplémentaire pour étudier les groupes eux-mêmes; 2) un but appliqué qui est illustré, par exemple, par sa brillante contribution à la cristallographie et à la mécanique quantique. Ni l'un ni l'autre de ces aspects n'est au fait reflété dans ce chapitre dont l'objectif est plus que modeste, à savoir: dire quelque chose de substantiel de la théorie des représentations, en partant uniquement de ce qui nous est connu en algèbre linéaire et en théorie des groupes.

§ 1. Définitions et exemples de représentations linéaires

1. Notions fondamentales.— A vrai dire, nous nous sommes déjà occupés de la théorie des représentations quand nous avons considéré (voir chap. 7, § 2) les opérations des groupes sur les ensembles. Prenons maintenant comme ensemble un espace vectoriel V de dimension n sur un corps commutatif K et considérons dans le groupe

$S(V)$ de toutes les transformations bijectives $V \mapsto V$ le sous-groupe $\mathrm{GL}(V)$ qui est le groupe des opérateurs linéaires inversibles sur V (ou le groupe des automorphismes de l'espace V). Il est clair que quel que soit le choix de la base $\{e_1, \ldots, e_n\}$ dans V, le groupe $\mathrm{GL}(V)$ devient un groupe ordinaire de matrices $\mathrm{GL}(n, K)$ qu'on peut considérer comme groupe des automorphismes de l'espace vectoriel $K^n = \mathbb{R}^n$. Dans ces conditions, à chaque opérateur linéaire $\mathcal{A} \in \mathrm{GL}(V)$ correspond une matrice $A = (a_{ij})$ telle que

$$\mathcal{A}e_j = \sum_{i=1}^{n} a_{ij}e_i; \quad a_{ij} \in K, \quad \det A \neq 0.$$

DÉFINITION 1.—*Soit G un groupe quelconque. Tout homomorphisme $\Phi\colon G \to \mathrm{GL}(V)$ s'appelle représentation linéaire du groupe G dans l'espace V.*

On dit qu'une représentation est exacte si son noyau $\mathrm{Ker}\,\Phi$ *est réduit au seul élément unité du groupe G, et triviale (ou représentation unité) si $\Phi(g) = \mathcal{E}$ est l'opérateur identique pour tous les éléments $g \in G$.*

La dimension $\dim_K V$ *est encore appelée degré de la représentation.*

Dans le cas où $K = \mathbb{Q}$, $\mathbb{R}$ ou $\mathbb{C}$ on parle respectivement de la représentation rationnelle, réelle ou complexe du groupe G.

Ainsi, la représentation linéaire est un couple (Φ, V) composé d'un *espace de représentation* V (ou d'un *G-espace*) et d'un homomorphisme $\Phi\colon G \to \mathrm{GL}(V)$. Par définition

$$\Phi(e) = \mathcal{E} \text{ est l'opérateur identique;}$$
$$\Phi(gh) = \Phi(g)\,\Phi(h), \text{ quels que soient } g, h \in G.$$

En convenant de désigner par $g * v$ l'action de l'opérateur linéaire $\Phi(g)$ sur un vecteur $v \in V$, nous obtenons les relations

$$\begin{aligned} g*(u+v) &= g*u+g*v, \quad & u, v \in V,\\ g*(\lambda v) &= \lambda(g*v), & \lambda \in K,\\ e*v &= v, &\\ (gh)*v &= g*(h*v), & \end{aligned} \tag{1}$$

qui simulent les propriétés des opérateurs linéaires, les deux dernières expressions remplaçant ce qui a été exprimé plus haut par le signe Φ (comparer avec (i), (ii) du § 2 du chap. 7). Les relations (1) mettent dans la représentation linéaire (Φ, V) à la première place le G-espace V, ce qui s'avère commode de faire pour une cause ou l'autre (par exemple, lorsque V n'est pas un espace vectoriel abstrait, mais l'une quelconque de ses réalisations concrètes).

D'autre part, l'espace V peut ne pas être mentionné si l'on entend par représentation linéaire tout simplement l'homomorphisme Φ du groupe G dans le groupe des matrices $\mathrm{GL}(n, K)$. Comme pré-

cédemment, $\Phi_{gh} = \Phi_g \Phi_h$, mais ici Φ_g est une matrice régulière et $\Phi_e = E$ est la matrice unité. L'interprétation matricielle est plus commode du point de vue du calcul, mais elle est moins invariante et est privée de toute interprétation spatiale. Au fait, il importe d'apprendre (ce qui n'est pas difficile) à passer librement des G-espaces aux représentations matricielles et inversement.

Rappelons à ce propos un fait bien connu en algèbre linéaire: deux matrices A, B, correspondant à un même opérateur linéaire dans des bases différentes, sont semblables, $B = CAC^{-1}$ (C est la matrice de passage d'une base à l'autre). Dans le cas des représentations, lorsqu'il s'agit du groupe des opérateurs linéaires, la dépendance vis-à-vis du choix de la base est prise en compte de la manière suivante.

DÉFINITION 2.— *On dit que deux représentations linéaires* (Φ, V), (Ψ, W) *d'un groupe* G *sont équivalentes (isomorphes ou semblables) s'il existe un isomorphisme des espaces vectoriels* $\sigma: V \to W$ *rendant commutatif le diagramme*

$$\begin{array}{ccc} V & \xrightarrow{\sigma} & W \\ {\scriptstyle \Phi(g)}\downarrow & & \downarrow{\scriptstyle \Psi(g)} \\ V & \xrightarrow{\sigma} & W \end{array}$$

pour tout $g \in G$, *c'est-à-dire si*

$$\Psi(g)\,\sigma = \sigma\Phi(g), \quad g \in G,$$

ou, ce qui revient au même,

$$\Psi(g) = \sigma\Phi(g)\,\sigma^{-1} \tag{2}$$

(comparer avec la définition de l'équivalence des opérations d'un groupe sur les ensembles, donnée au chap. 7, § 2, exercice 1).

Nous écrirons parfois $\Phi \approx \Psi$ pour des représentations équivalentes et $\Phi \not\approx \Psi$ pour des représentations non équivalentes.

Donnons encore deux variantes de la définition 2.

a) **Langage de G-espaces.**—Soient G un groupe et V: $(g, v) \mapsto g * v$, W: $(g, w) \mapsto g \square w$ deux G-espaces munis des opérations $*$, $\square$ qui satisfont aux conditions (1). L'isomorphisme $\sigma: V \to W$ des espaces vectoriels est dit *isomorphisme des G-espaces* si

$$g \square \sigma(v) = \sigma(g * v) \tag{2'}$$

pour tout $g \in G$ et tout $v \in V$. On dit encore que l'application σ est permutable avec l'opération de G.

b) **Langage matriciel.**— Si $V = \langle v_1, \ldots, v_n \rangle$, $W = \langle w_1, \ldots, w_n \rangle$ et Φ_g, Ψ_g sont des matrices des opérateurs linéaires $\Phi(g)$, $\Psi(g)$ par rapport aux bases choisies, la condition d'équiva-

lence (2) s'écrit sous la forme

$$\Psi_g = C\Phi_g C^{-1}, \tag{2''}$$

où C est une matrice régulière, la même pour tous les $g \in G$. Les coefficients de toutes les matrices considérées appartiennent au même corps K.

La relation de similitude des matrices exprimée par la condition (2″) est une relation d'équivalence qui partage l'ensemble $M_n(K)$ en classes disjointes. Respectivement, les représentations du groupe G sont partagées en classes de représentations équivalentes. Nous verrons plus loin que ce sont justement les classes de représentations équivalentes qui sont intéressantes et substantielles pour la théorie des représentations.

En nous reportant de nouveau au cours d'Algèbre linéaire, proposons-nous de représenter d'une manière plus concrète l'opération du groupe $\Phi(G)$ sur l'espace V. Il peut exister dans V un sous-espace U invariant par rapport à l'opérateur linéaire $\mathcal{A}: V \to V$, si bien que $u \in U \Rightarrow \mathcal{A}u \in U$. En complétant une base arbitraire $\{e_1, \ldots, e_k\}$ de U jusqu'à la base de tout l'espace $V = \langle e_1, \ldots, e_k, e_{k+1}, \ldots, e_n \rangle$, nous verrons que la matrice de l'opérateur $\mathcal{A}$ prendra dans la base $\{e_1, \ldots, e_n\}$ une forme triangulaire par blocs

$$A = \left\| \begin{matrix} A_1 & A_0 \\ 0 & A_2 \end{matrix} \right\|.$$

Le bloc A_1 correspond au sous-espace invariant U, et le bloc A_2 à l'espace quotient V/U. Si A_0 est la matrice nulle, alors $A = A_1 \dotplus A_2$ est la somme directe des blocs, et $V = U \oplus W$ la somme directe des sous-espaces invariants.

L'existence d'un sous-espace invariant propre par rapport à $\mathcal{A}$ est toujours assurée si le corps de base K est algébriquement clos et $\dim V > 1$ (voir chap. 6, § 3). Si, par exemple, $K = \mathbb{C}$ est le corps des nombres complexes, il existe un vecteur $v \in V$, $v \neq 0$, tel que $\mathcal{A}v = \lambda v$. Ici, λ est racine du polynôme caractéristique

$$f_{\mathcal{A}}(t) = |tE - A| = t^n - (\operatorname{tr} A)\, t^{n-1} + \ldots + (-1)^n \det A$$

(A est une matrice arbitraire de l'opérateur linéaire $\mathcal{A}$). Cette circonstance permet de choisir dans V une base par rapport à laquelle la matrice A prend une forme triangulaire

$$A = \left\| \begin{matrix} \lambda_1 & & & * \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{matrix} \right\|$$

avec les racines caractéristiques $\lambda_1, \lambda_2, \ldots, \lambda_n$ sur la diagonale. Une analyse un peu plus fine se termine par la réduction de A à la *forme réduite de Jordan* $J(A)$ (voir Annexe), c'est-à-dire à la somme directe des *matrices de Jordan*

$$J_{m,\lambda}=\left\|\begin{matrix}\lambda & 1 & 0 & \ldots & 0\\ 0 & \lambda & 1 & \ldots & 0\\ \cdot & \cdot & \cdot & \cdot\cdot\cdot & \cdot\\ 0 & 0 & 0 & \ldots & \lambda\end{matrix}\right\|$$

(m est l'ordre de la matrice de Jordan, λ est l'une des racines caractéristiques).

Remarquons que, si $A^q = E$, alors $J^q_{m,\lambda} = E_m$ est la matrice unité $m \times m$ pour chaque matrice de Jordan $J_{m,\lambda}$ de la matrice A, ce qui ne peut évidemment avoir lieu que si $m = 1$ et λ est la racine q-ième de l'unité (nous supposons toujours que $K = \mathbb{C}$). Par suite,

$$A^q=E \Rightarrow CAC^{-1}=\left\|\begin{matrix}\lambda_1 & & & 0\\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n\end{matrix}\right\|, \quad \lambda_i^q=1, \tag{3}$$

pour une certaine matrice inversible C.

Toutes les considérations, relatives à un opérateur linéaire isolé $\mathcal{A}: V \to V$, seront bien utiles lorsque nous passerons au groupe $\Phi(g)$, $g \in G$, des opérateurs linéaires.

Définition 3.— *Soit (Φ, V) une représentation linéaire d'un groupe G. Le sous-espace $U \subset V$ est dit invariant (ou stable) par G, si $\Phi(g)\,u \in U$ pour tous les $u \in U$ et tous les $g \in G$.*

Le sous-espace nul et l'espace V lui-même de la représentation Φ sont des sous-espaces invariants triviaux.

Si une représentation ne possède que des sous-espaces invariants triviaux, on dit qu'elle est irréductible. Une représentation est dite réductible si elle admet au moins un sous-espace invariant non trivial.

Il résulte de tout ce qui précède que dans le cas d'une représentation réductible (Φ, V), avec un sous-espace invariant U, l'espace V possède une base pour laquelle

$$\Phi_g=\left\|\begin{matrix}\Phi'_g & \Phi^0_g\\ 0 & \Phi''_g\end{matrix}\right\| \tag{4}$$

pour tous les $g \in G$. Puisque $\Phi'_{gh} = \Phi'_g\Phi'_h, \Phi'_e = E_k$ et $\Phi'_g(U) \subset$ $\subset U$, l'application $\Phi': g \mapsto \Phi'_g$ définit une représentation sur U que l'on appelle *sous-représentation* de Φ. Une représentation est

aussi définie sur l'espace quotient V/U. Elle est appelée *représentation quotient* et est définie par les matrices Φ_g'', $g \in G$.

Si la base dans V peut être choisie de manière que toutes les matrices Φ_g^0 figurant dans (4) soient nulles, on dit que l'on a affaire à une représentation *décomposable* Φ ou plus exactement à une *somme directe des représentations* $\Phi = \Phi' \dot{+} \Phi''$. La décomposition de (Φ, V) en une somme directe est réalisable si, et seulement si, le sous-espace invariant $U \subset V$ possède un sous-espace invariant *supplémentaire* W tel que $V = U \oplus W$ soit une décomposition en somme directe des sous-espaces, et $\Phi(U) \subset U$, $\Phi(W) \subset W$. S'il en est ainsi, $\Phi' = \Phi|_U$, $\Phi'' = \Phi|_W$ sont des restrictions de Φ respectivement à U et à W.

Une représentation linéaire (Φ, V) est dite *indécomposable* si elle ne peut pas être représentée sous la forme d'une somme directe de deux sous-représentations non triviales. On parle aussi d'un *G-espace V indécomposable.*

En scindant successivement, si cela est possible, V, U, W, etc., en sommes directes des sous-espaces invariants, nous obtiendrons une somme directe $V = V_1 \oplus \ldots \oplus V_r$ de plusieurs sous-espaces invariants (respectivement une somme directe $\Phi = \Phi^{(1)} \dot{+} \ldots \dot{+} \Phi^{(r)}$ de plusieurs représentations). Avec un choix convenable de la base dans V, les matrices des opérateurs linéaires prendront la forme

$$\Phi_g = \begin{Vmatrix} \Phi_g^{(1)} & 0 & \ldots & 0 \\ 0 & \Phi_g^{(2)} & \ldots & 0 \\ \cdot & \cdot & \cdot\ \cdot\ \cdot & \cdot \\ 0 & 0 & \ldots & \Phi_g^{(r)} \end{Vmatrix}.$$

DÉFINITION 4. — *Une représentation linéaire* (Φ, V) *d'un groupe* G *est dite complètement réductible si elle est une somme directe des représentations irréductibles. La même terminologie s'applique aux G-espaces.*

Par intuition, il est clair que les représentations irréductibles jouent le rôle de blocs de construction, à partir desquels on forme des représentations linéaires arbitraires. Les représentations complètement réductibles sont obtenues en utilisant la construction la plus simple, c'est-à-dire la somme directe. Nous verrons par la suite que dans beaucoup de cas cela suffit pour décrire toutes les représentations. Signalons que certains groupes ayant une grande importance pour la physique, comme par exemple les groupes de Lorentz, comportent des *représentations irréductibles de dimension infinie.* On comprend que ces dernières ne se réduisent aucunement à des représentations de dimension finie et doivent donc être étudiées séparément.

2. Exemples de représentations linéaires.— Nous avons introduit toutes les définitions fondamentales de la théorie des représentations. Il ne reste qu'à les remplir d'un contenu réel. A cet effet, il est utile de se familiariser tout d'abord avec les exemples que nous donnons ci-dessous et de les bien comprendre.

EXEMPLE 1. —Le groupe linéaire complet GL (n, K) sur un corps commutatif K admet, par définition, une représentation linéaire irréductible et exacte de degré n, avec un espace de représentation $V = K^n$. Sur le même espace, tout groupe linéaire $H \subset \mathrm{GL}(n, K)$ opère exactement mais peut-être de façon réductible.

Des remarques analogues sont applicables à d'autres groupes classiques indiqués au chapitre 7, § 1. Par exemple, le groupe unitaire U (n) opère de façon irréductible sur l'espace hermitien, et le groupe orthogonal O (n) fait de même sur l'espace euclidien. Cela résulte directement d'une proposition plus forte, démontrée en Algèbre linéaire, d'après laquelle les groupes U (n) et O (n) opèrent transitivement (au sens de l'exemple 3 du chap. 7, § 2, n° 3) sur l'ensemble des vecteurs unitaires.

EXEMPLE 2.— En faisant opérer le groupe GL (n, K) sur l'espace vectoriel $M_n(K)$ des matrices d'ordre n suivant la règle $\Psi_A : X \mapsto AX$ ($A \in \mathrm{GL}(n, K)$, $X \in M_n(K)$), nous nous assurons sans difficulté que $\Psi_A(\alpha X + \beta Y) = \alpha\Psi_A X + \beta\Psi_A Y$ et $\Psi_{AB} = \Psi_A\Psi_B$. Par conséquent, $(\Psi, M_n(K))$ est une représentation linéaire de degré n^2. Soit $M_n^{(i)}(K)$ le sous-espace des matrices

$$\begin{Vmatrix} 0 & \dots & x_{1i} & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \dots & x_{ni} & \dots & 0 \end{Vmatrix}$$

avec une seule colonne non nulle $X^{(i)}$. Il est facile de vérifier que ce sous-espace est invariant par Ψ_A, $A \in \mathrm{GL}(n, K)$, irréductible et isomorphe (en tant que GL(n, K)-espace) à l'espace K^n sur lequel opère le groupe GL (n, K). Ainsi

$$M_n(K) = M_n^{(1)}(K) \oplus \dots \oplus M_n^{(n)}(K)$$

est une décomposition en somme directe de n GL (n, k)-sous-espaces isomorphes, à laquelle correspond la décomposition

$$\Psi = \Psi^{(1)} \dot{+} \dots \dot{+} \Psi^{(n)}$$

en somme directe de n représentations équivalentes. Symboliquement, ce fait s'écrit sous la forme

$$M_n(K) \cong nM_n^{(1)}(K); \quad \Psi \approx n\Psi^{(1)}.$$

EXEMPLE 3.— Définissons maintenant l'opération Φ du groupe GL (n, K) sur $M_n(K)$ en posant $\Phi_A : X \mapsto AXA^{-1}$. De nouveau $(\Phi, M_n(K))$ est une représentation linéaire de degré n^2. Si $X = (x_{ij})$, alors $\mathrm{tr}X = \sum_{i=1}^{n} x_{ii}$ est comme à l'ordinaire la trace de la matrice X. Il est bien connu que $\mathrm{tr}(\alpha X + \beta Y) = \alpha\,\mathrm{tr}\,X + \beta\,\mathrm{tr}\,Y$ (linéarité de la fonction tr) et $\mathrm{tr}\,\Phi_A(X) = \mathrm{tr}\,X$. Il en résulte que l'ensemble $M_n^0(K)$ des matrices à trace nulle est un sous-espace invariant par Φ. D'autre part, $\Phi_A(\lambda E) = \lambda E$ et $\mathrm{tr}\,\lambda E = n\lambda$. Ainsi, dans le cas d'un corps K de caractéristique nulle, on a une décomposition en somme directe des GL (n, K)-sous-espaces

$$M_n(K) = \langle E \rangle \oplus M_n^0(K) \tag{5}$$

de dimensions 1 et $n^2 - 1$ respectivement. Remarquons que pour $n = p$ et $K = Z_p$, la décomposition de la forme (5), n'existe pas, car dans ce cas tr $E = 0$.

D'après la définition, la forme réduite de Jordan $J(X)$ de la matrice X n'est rien d'autre que le représentant le plus simple et commode de la GL $(n, \mathbb{C})$-orbite contenant X. La restriction de Φ à tout sous-groupe $H \subset$ GL (n, K) rend naturelle la question relative aux représentants canoniques des H-orbites.

EXEMPLE 4.—Dans l'exemple précédent, posons $K = \mathbb{R}$ et considérons la restriction de Φ au groupe orthogonal O (n). Puisque $A \in$ O $(n) \Longleftrightarrow {}^tA = A^{-1}$, on a ${}^t(AXA^{-1}) = {}^tA^{-1} \cdot {}^tX \cdot {}^tA = \varepsilon AXA^{-1}$ pour ${}^tX = \varepsilon X$, $\varepsilon = \pm 1$. Par conséquent, l'espace de représentation $M_n(\mathbb{R})$ du groupe O (n) s'écrit sous la forme d'une somme de O (n)-sous-espaces

$$M_n(\mathbb{R}) = \langle E\rangle_{\mathbb{R}} \oplus M_n^+(\mathbb{R}) \oplus M_n^-(\mathbb{R}),$$

c'est-à-dire de l'espace de dimension un $\langle E\rangle_{\mathbb{R}}$ des matrices scalaires, de l'espace de dimension $(n+2)(n-1)/2$ des matrices symétriques de trace nulle et de l'espace de dimension $n(n-1)/2$ des matrices antisymétriques. Il est bien connu qu'il existe une correspondance biunivoque entre les matrices symétriques (antisymétriques) et les formes bilinéaires symétriques (respectivement antisymétriques). L'opération de O (n) sur $\langle E\rangle_{\mathbb{R}} \oplus M_n^+(\mathbb{R})$ et sur $M_n^-(\mathbb{R})$ est transférée sur les espaces des formes correspondantes. Le théorème sur la réduction de la forme quadratique $q(x)$ aux axes principaux n'est rien d'autre que la possibilité de choix, dans la O (n)-orbite contenant $q(x)$, de la forme diagonale $\sum \lambda_i x_i^2$: avec λ_i réels définis de façon unique à une permutation près.

En remplaçant $\mathbb{R}$ par $\mathbb{C}$ et O (n) par le groupe unitaire U (n), nous obtenons la décomposition

$$M_n(\mathbb{C}) = \langle E\rangle\mathbb{C} \oplus M_n^+(\mathbb{C}) \oplus M_n^-(\mathbb{C})$$

en somme directe de U (n)-sous-espaces des matrices scalaires hermitiennes à trace nulle et hermitiennes gauches.

EXEMPLE 5.— Soit G un groupe de permutations, opérant sur un certain ensemble Ω dont le nombre d'éléments $|\Omega| = n > 1$, c'est-à-dire $G \subset S_n$. L'espace vectoriel

$$V = \langle e_i \mid i \in \Omega\rangle_K$$

sur un corps commutatif K de caractéristique nulle et de base numérotée par les éléments de l'ensemble Ω sera transformé en un G-espace si l'on pose

$$\Phi(g)\Big(\sum_{i\in\Omega} \lambda_i e_i\Big) = \sum_{i\in\Omega} \lambda_i \Phi(g)\, e_i = \sum_{i\in\Omega} \lambda_i e_{g(i)}$$

($i \longmapsto g(i)$ est l'opération de la permutation $g \in G$ sur $i \in \Omega$). Puisque $(gh)(i) = g(h(i))$, on obtient une représentation linéaire de degré n du groupe G. Elle n'est jamais irréductible, car

$$V = \Big\langle \sum_{i\in\Omega} e_i \Big\rangle \oplus \Big\{ \sum_{\lambda_1+\ldots+\lambda_n=0} \lambda_i e_i \mid \lambda_i \in K \Big\} \tag{6}$$

est la décomposition en somme directe des sous-espaces invariants de dimensions 1 et $(n-1)$ (si car $k = p > 0$ et $p \mid n$, la somme directe ne s'obtient plus).

Considérons deux cas particuliers.

a) $G = S_n$.— Le monomorphisme $S_n \to$ GL $(n, \mathbb{R})$ construit au chapitre 4, § 3, n° 5, coïncide avec notre représentation linéaire Φ si l'on prend pour e_i la i-ième colonne $E^{(i)}$. La décomposition (6) exprime que pour S_n il existe une injection plus économique $S_n \to$ GL $(n-1, \mathbb{R})$. Plus tard, nous démontrerons

l'irréductibilité de cette représentation linéaire de degré $n - 1$ (même sur le corps $\mathbb{C}$).

b) *Représentation régulière.*— Soit G un groupe fini arbitraire. En posant $\Omega = G$, nous obtiendrons un G-espace $V = \langle e_g \mid g \in G \rangle$ dit régulier et respectivement une *représentation régulière* (ρ, V) du groupe G: $\rho(a) e_g = e_{ag}$, quels que soient $a, g \in G$. A vrai dire, nous avons déjà rencontré une représentation régulière, avec des notations légèrement différentes, lors de la démonstration du théorème de Cayley (chap. 4, § 3), mais nous nous sommes intéressés alors non pas à l'espace V mais à l'ensemble $\{e_g\}$ de ses vecteurs de base. L'importance de la représentation régulière d'un groupe fini G consiste en ce qu'elle contient toutes les représentations irréductibles de G, considérées à une équivalence près (voir plus loin § 5).

Exemple 6.— Une représentation de degré 1 est tout simplement un homomorphisme $\Phi: G \to K^*$ du groupe G dans le groupe multiplicatif du corps commutatif K (K est un espace vectoriel de dimension un sur lui-même). Le groupe multiplicatif du corps commutatif étant abélien, $\mathrm{Ker}\,\Phi \supset G'$, où G' est le sous-groupe dérivé du groupe G (chap. 7, § 3, théorème 4). Remarquons que l'*équivalence de deux représentations de dimension un* Φ', Φ'' (avec le même espace de représentation) *signifie leur coïncidence*, car $a\Phi'(g)a^{-1} = \Phi''(g) \Rightarrow \Phi'(g) = \Phi''(g) \Rightarrow \Phi' = \Phi''$. Soit $g^n = e$. Alors $\Phi(g)^n = \Phi(g^n) = \Phi(e) = 1$, c'est-à-dire $\Phi(g)$ est une racine de l'unité. Le noyau de toute représentation de dimension un peut être non trivial même pour un groupe cyclique G. Si par exemple $G = Z_4$ et $K = Z_7$, alors $\mathrm{Ker}\,\Phi \supset 2Z_4$. D'autre part, dans le cas où $K = \mathbb{C}$, tout groupe cyclique possède une représentation exacte de dimension un.

a) $G = (\mathbb{Z}, +)$.— La représentation $k \mapsto \lambda^k$ est exacte pour $|\lambda| \neq 1$. Si $|\lambda| = 1$, la formule d'Euler donne $\lambda = e^{2\pi i\theta}$, $\theta \in \mathbb{R}$, et le noyau de l'application $k \mapsto e^{2\pi i\theta k}$ n'est différent de zéro que pour $\theta \in \mathbb{Q}$.

Le groupe $\mathbb{Z}$ possède des représentations complexes indécomposables de degré aussi élevé que l'on veut, qui ne sont pas pourtant irréductibles. Il suffit de se rappeler le théorème sur la forme réduite de Jordan de la matrice et de considérer l'application

$$k \mapsto J^k_{m,1} = \begin{Vmatrix} 1 & 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & 1 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & 1 & 1 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{Vmatrix}^k .$$

b) $G = \langle a \mid a^n = e \rangle$.— Soit $\varepsilon = e^{\frac{2\pi i}{n}}$ une racine primitive n-ième de l'unité. Parmi les n représentations de dimension un

$$\Phi^{(m)}: a^k \mapsto \varepsilon^{mk}, \quad m = 0, 1, \dots, n-1, \tag{7}$$

il y a $\varphi(n)$ qui sont exactes. Signalons un fait intéressant: *un groupe cyclique d'ordre n possède exactement n représentations deux à deux non équivalentes irréductibles sur $\mathbb{C}$. Toutes ces représentations sont de dimension un et de la forme* (7).

En effet, il suffit de s'assurer que le groupe cyclique fini ne possède pas de représentations irréductibles sur $\mathbb{C}$ de dimension >1. Or, avant de donner la définition 3, nous avons signalé le fait que tout opérateur linéaire $\Phi(g)$ d'ordre fini est diagonalisable sur $\mathbb{C}$. Dans le cas considéré, cela est équivalent à ce que la représentation Φ est complètement réductible. *Si* $\dim \Phi = r$, *alors Φ se décompose en une somme directe de r représentations à une dimension.* ∎

Pour le groupe cyclique d'ordre fini, nous avons obtenu au fait la description de toutes les représentations complexes linéaires. A une équivalence près,

on a

$$\Phi_g = \begin{Vmatrix} \Phi_g^{(i_1)} & & 0 \\ & \ddots & \\ 0 & & \Phi_g^{(i_r)} \end{Vmatrix},$$

où $\Phi^{(m)}$ est l'une des représentations de la forme (7).

Notre but est d'établir de pareilles lois dans le cas général.

EXEMPLE 7.— Les exemples qui précèdent montrent déjà que les propriétés de la représentation linéaire Φ du groupe G dépendent beaucoup du corps de base K. Apportons plus de clarté à cette question.

Un groupe cyclique $G = \langle a \mid a^p = e\rangle$ d'ordre un nombre premier p, opérant sur un espace vectoriel de $V = \langle v_1, v_2\rangle$ de dimension deux sur un corps commutatif arbitraire K de caractéristique p, suivant la loi $a * v_1 = v_1$, $a * v_2 = v_1 + v_2$, définit une représentation *indécomposable* (Φ, V)

$$a^k \mapsto \Phi_a^k \begin{Vmatrix} 1 & k \\ 0 & 1 \end{Vmatrix}, \quad 0 \leqslant k \leqslant p-1.$$

En effet, la matrice Φ_a possède une racine caractéristique 1 de multiplicité 2. De ce fait, la possibilité de décomposer Φ en somme directe de deux représentations de dimension un signifierait l'existence d'une matrice inversible C pour laquelle $C\Phi_a C^{-1} = \begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix} = E$. On aurait alors $\Phi_a = C^{-1}EC = E$, ce qui n'est pas vrai.

Soit ensuite $G = \langle a \mid a^3 = e\rangle$ un groupe cyclique d'ordre 3 et soit $K = \mathbb{R}$. La représentation de dimension deux (Φ, V), avec $V = \langle v_1, v_2\rangle$, définie dans la base donnée par la matrice

$$\Phi_a = \begin{Vmatrix} -1 & -1 \\ 1 & 0 \end{Vmatrix},$$

est irréductible, car le polynôme caractéristique $t^2 + t + 1$ de cette matrice n'a pas de racines réelles. Or, si V est considéré sur $\mathbb{C}$, il se décompose naturellement en somme directe des G-sous-espaces de dimension un

$$V = \langle v_1, + \varepsilon^{-1} v_2\rangle \oplus \langle v_1 + \varepsilon v_2\rangle$$

et

$$C\Phi_a C^{-1} = \begin{Vmatrix} \varepsilon & 0 \\ 0 & \varepsilon^{-1} \end{Vmatrix}; \quad \varepsilon = \frac{-1+\sqrt{-3}}{2}, \quad C = \begin{Vmatrix} 1 & -\varepsilon^{-1} \\ 1 & -\varepsilon \end{Vmatrix}.$$

On voit donc que l'extension du corps peut faire perdre à la représentation sa propriété d'être irréductible.

Dans la suite de cet ouvrage, le corps de base K sera présenté, à de rares exceptions, par le corps des nombres complexes (le plus important au point de vue pratique) ou par un corps arbitraire algébriquement clos de caractéristique nulle.

EXERCICES

1. Le groupe SO (2) est défini par sa représentation naturelle de dimension deux

$$\Phi'(\theta) = \begin{Vmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{Vmatrix},$$

irréductible sur $\mathbb{R}$. Vérifier que

$$A\Phi'(\theta)A^{-1} = \begin{Vmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{Vmatrix} \quad \text{pour} \quad A = \frac{1}{\sqrt{2}} \begin{Vmatrix} 1 & i \\ i & 1 \end{Vmatrix} \in \mathrm{GL}(2, \mathbb{C}).$$

Par conséquent, Φ' est une somme directe de deux représentations de dimension un, non équivalentes (dans le cas considéré, tout simplement différentes).

2. Le GL $(n, \mathbb{C})$-espace $M_n^0(\mathbb{C})$ dans la décomposition (5) est-il irréductible pour $n = 2$ et 3? (Réponse: oui.)

3. Soient Φ et Ψ deux représentations complexes irréductibles du groupe cyclique $\langle a \mid a^n = e \rangle$ d'ordre n. Montrer que

$$\frac{1}{n} \sum_{k=0}^{n-1} \Phi(a^k)\overline{\Psi(a^k)} = \begin{cases} 1 & \text{si} \quad \Phi \approx \Psi, \\ 0 & \text{si} \quad \Phi \not\approx \Psi. \end{cases}$$

4. En s'appuyant sur l'exercice 3, s'assurer que l'assertion suivante est vraie. Toute fonction f à valeurs complexes sur un groupe cyclique fini $\langle a \mid a^n = e \rangle$ peut s'écrire sous la forme d'une décomposition suivant les « harmoniques élémentaires »

$$f(a^k) = \sum_{m=0}^{n-1} c_m \varepsilon^{mk}, \qquad \varepsilon = e^{\frac{2\pi i}{n}}.$$

Les « coefficients de Fourier » c_m se calculent par la formule

$$c_m = \frac{1}{n} \sum_{k=0}^{n-1} f(a^k)\, \varepsilon^{-mk}.$$

5. De la formule donnant le nombre de colliers (voir début du chapitre) déduire les conséquences élémentaires: a) $q^p - q \equiv 0 \pmod{p}$ (théorème de Fermat; voir chap. 4, § 4); b) $\sum_{d \mid n} \varphi(d) = n$.

§ 2. Représentations unitaires et réductibles

1. Représentations unitaires. — Rappelons que, dans le cours d'Algèbre linéaire, une forme non dégénérée $(u, v) \mapsto (u \mid v)$ sur un espace vectoriel V sur $\mathbb{C}$ est dite *hermitienne* si

$$\begin{gathered} (u \mid v) = \overline{(v \mid u)}, \\ (\alpha u + \beta v \mid w) = \alpha\,(u \mid w) + \beta\,(v \mid w), \\ (v \mid v) > 0 \text{ pour tout } v \neq 0 \end{gathered} \tag{1}$$

(comme toujours, $z \mapsto \bar{z}$ est un automorphisme de conjugaison complexe). L'espace V muni d'une forme hermitienne non dégénérée $(u \mid v)$ s'appelle espace *hermitien*. Son analogue réel est un espace euclidien muni d'un produit scalaire défini par une forme bilinéaire symétrique non dégénérée. En prenant dans V une base $e_1, \ldots, e_n$, nous écrirons la forme $(u \mid v)$ pour $u = \sum u_i e_i$, $v = \sum v_j e_j$ de la manière suivante:

$$(u \mid v) = \sum h_{ij} u_i \bar{v}_j.$$

La matrice $H = (h_{ij})$ satisfait à la condition $\bar{h}_{ij} = h_{ji}$ et s'appelle aussi matrice *hermitienne*. Nous avons déjà utilisé une telle terminologie au chapitre 7, § 1.

Il existe une base orthonormée définie par la condition $(e_i \mid e_j) = \delta_{ij}$ par rapport à laquelle

$$(u \mid v) = \sum_{i=1}^{n} u_i \bar{v}_i.$$

Un opérateur linéaire $\mathcal{A}: V \to V$ qui conserve cette forme, c'est-à-dire possède la propriété $(\mathcal{A}u/\mathcal{A}v) = (u \mid v)$, s'appelle opérateur *unitaire*. Son analogue réel est un opérateur *orthogonal*. La condition qu'un opérateur est unitaire, écrite sous forme matricielle $A \cdot {}^t\overline{A} = E$, avec $A = (a_{ij})$, ${}^t\overline{A} = A^* = (\overline{a}_{ji})$, a été déjà rencontrée au chapitre 7. En désignant par $\mathcal{A}^*$ l'opérateur linéaire de matrice ${}^t\overline{A} = A^*$, nous pouvons exprimer cette condition sous la forme $\mathcal{A} \cdot \mathcal{A}^* = \mathcal{E} = \mathcal{A}^* \cdot \mathcal{A}$.

On convient de désigner le groupe de toutes les matrices unitaires (groupe des opérateurs unitaires ou tout simplement groupe unitaire) par le symbole U (n) que nous connaissons. Par définition, U $(n) \subset$ GL $(n, \mathbb{C})$, *et si la représentation* $\Phi: G \to$ GL $(n, \mathbb{C})$ *est telle que* Im $\Phi \subset$ U (n), *alors on dit que* (Φ, V) *est une représentation unitaire.*

THÉORÈME 1. — *Toute représentation linéaire* (Φ, V) *sur* $\mathbb{C}$ *d'un groupe fini* G *est équivalente à une représentation unitaire.*

DÉMONSTRATION. —Choisissons dans l'espace de représentation V du groupe G une forme hermitienne non dégénérée quelconque $H: (u, v) \mapsto H(u, v) = \sum h_{ij} u_i \bar{v}_j$ (écriture relative à une certaine base $f_1, \ldots, f_n$ de l'espace V) et considérons la forme $(u \mid v)$ obtenue à partir de $H(u, v)$ en faisant une « moyenne » par rapport à G:

$$(u \mid v) = |G|^{-1} \sum_{g \in G} H(\Phi(g) u, \Phi(g) v). \tag{2}$$

Le facteur $|G|^{-1}$ est sans importance, on ne le met que pour obtenir, dans le cas où H est unitaire, l'égalité $(u \mid v) = H(u, v)$. Puisque

$$H(\Phi(g) u, \Phi(g) v) = \overline{H(\Phi(g) v, \Phi(g) u)},$$

$$H(\Phi(g)(\alpha u + \beta v), \Phi(g) w) =$$
$$= H(\alpha\Phi(g) u + \beta\Phi(g) v, \Phi(g) w) =$$
$$= \alpha H(\Phi(g) u, \Phi(g) w) + \beta H(\Phi(g) v, \Phi(g) w),$$
$$H(\Phi(g) v, \Phi(g) v) > 0$$

pour $v \neq 0$ et tout $g \in G$, la forme (2) satisfait aux conditions (1) et est donc une forme hermitienne non dégénérée.

De plus (ce qui est l'essentiel)

$$(\Phi(g)u \mid \Phi(g)v) =$$
$$= |G|^{-1} \sum_{h\in G} H(\Phi(g)\Phi(h)u, \Phi(g)\Phi(h)v) =$$
$$= |G|^{-1} \sum_{h\in G} H(\Phi(gh)u, \Phi(gh)v) =$$
$$= |G|^{-1} \sum_{t\in G} H(\Phi(t)u, \Phi(t)v) = (u \mid v),$$

c'est-à-dire l'opérateur $\Phi(g)$ laisse invariante la forme $(u \mid v)$ quel que soit $g \in G$. Choisissons dans V une base $e_1, \ldots, e_n$ orthonormée par rapport à la forme $(u \mid v)$. Alors, dans cette base, les matrices Φ_g des opérateurs $\Phi(g)$ seront unitaires.

REMARQUES. 1) L'assertion du théorème 1 ne découle pas automatiquement du fait connu que chaque matrice Φ_g, avec $g^m = e$, est semblable à la matrice unitaire diag $\{\lambda_1, \ldots, \lambda_n\}$, avec $\lambda_i^m = 1$.

2) Dans le cas réel, le raisonnement tout à fait analogue montre que la représentation (Φ, V) est équivalente à une représentation orthogonale.

3) Pour beaucoup de raisons, les représentations unitaires jouent un rôle important dans les applications de la théorie des représentations, et il est bien remarquable que le théorème 1 reste vrai pour une classe sensiblement plus large de groupes compacts tels que U (n) et O (n). La démonstration est la même, mais la sommation sur les éléments du groupe est remplacée par l'intégration (suivant une certaine mesure) sur le groupe. Rappelons que le groupe compact SU (2) est géométriquement indiscernable de la sphère S^3 de dimension trois, et on peut donc parler de son volume, par exemple. En général, il existe un parallélisme bien marqué dans la théorie des représentations des groupes finis et des groupes compacts, mais son exposé sortirait nettement du cadre du présent ouvrage. L'exemple 6, a) du § 1 montre que les représentations des groupes non compacts (par exemple, $G = \mathbb{Z}$) ne doivent pas être nécessairement unitaires.

Notons, avant de clore ce numéro, que bien que la démonstration du théorème 1 soit constructive, il serait peu pratique de l'utiliser pour la recherche de la réalisation unitaire d'une représentation donnée. Par exemple, pour un groupe G engendré par des éléments $a_1, \ldots, a_d$ il suffit d'assurer que les matrices $\Phi_{a_1}, \ldots, \Phi_{a_d}$ soient unitaires. Alors, le groupe $\langle \Phi_{a_1}, \ldots, \Phi_{a_d} \rangle = \Phi(G)$ sera, lui aussi, unitaire.

EXEMPLE 1.— Le groupe symétrique $S_3 = \langle (12), (123) \rangle$ possède une représentation Φ de dimension deux, qui est contenue comme terme direct dans la représentation naturelle de dimension trois (voir exemple 5 du § 1). A savoir, si $\Phi(\pi)e_i = e_{\pi(i)}$, $i = 1, 2, 3$, et $f_1 = e_1 - e_3$, $f_2 = e_2 - e_3$, alors

$$\Phi((12))f_1 = e_2 - e_3 = f_2, \qquad \Phi((12))f_2 = e_1 - e_3 = f_1,$$
$$\Phi((123))f_1 = e_2 - e_1 = -f_1 + f_2, \quad \Phi((123))f_2 = e_3 - e_1 = -f_1.$$

Puisque $\pi = (123)^i(12)^j$, où $i = 0$, 1 ou 2, et $j = 0$ ou 1, on obtient sans peine toutes les matrices

$$e \mapsto \begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix}, \quad (12) \mapsto \begin{Vmatrix} 0 & 1 \\ 1 & 0 \end{Vmatrix}, \quad (13) \mapsto \begin{Vmatrix} -1 & -1 \\ 0 & 1 \end{Vmatrix},$$

$$(23) \mapsto \begin{Vmatrix} 1 & 0 \\ -1 & -1 \end{Vmatrix}, \quad (123) \mapsto \begin{Vmatrix} -1 & -1 \\ 1 & 0 \end{Vmatrix}, \quad (132) \mapsto \begin{Vmatrix} 0 & 1 \\ -1 & -1 \end{Vmatrix}.$$

Des relations $\det \begin{Vmatrix} -1 & -1 \\ 1 & 0 \end{Vmatrix} = 1$ et $(123)^3 = e$ on déduit que pour une matrice régulière C on a

$$C \begin{Vmatrix} -1 & -1 \\ 1 & 0 \end{Vmatrix} C^{-1} = \begin{Vmatrix} \varepsilon & 0 \\ 0 & \varepsilon^{-1} \end{Vmatrix}, \quad \varepsilon = \frac{-1+\sqrt{-3}}{2}.$$

La conjugaison à l'aide de C doit garder la propriété de la matrice $\begin{Vmatrix} 0 & 1 \\ 1 & 0 \end{Vmatrix}$ d'être unitaire. Les conditions linéaires

$$C \begin{Vmatrix} 0 & 1 \\ 1 & 0 \end{Vmatrix} = \begin{Vmatrix} 0 & 1 \\ 1 & 0 \end{Vmatrix} C, \quad C \begin{Vmatrix} -1 & -1 \\ 1 & 0 \end{Vmatrix} = \begin{Vmatrix} \varepsilon & 0 \\ 0 & \varepsilon^{-1} \end{Vmatrix} C, \quad C = \begin{Vmatrix} \alpha & \beta \\ \gamma & \delta \end{Vmatrix}$$

sont vérifiées par la matrice

$$C = \begin{Vmatrix} 1 & -\varepsilon^2 \\ -\varepsilon^2 & 1 \end{Vmatrix}.$$

Maintenant nous sommes en mesure d'écrire les représentations unitaire du groupe S_3 que nous connaissons: la représentation unitaire $\Phi^{(1)}$, $\Phi^{(2)}: \pi \mapsto \mapsto \operatorname{sgn}(\pi) = \pm 1$ et la représentation $\Phi^{(3)} \approx \Psi$ de dimension deux que nous venons d'obtenir. Pour faire des références dans la suite, il est commode de donner la table suivante:

Φ \ g	e	(12)	(13)	(23)	(123)	(132)
$\Phi^{(1)}$	1	1	1	1	1	1
$\Phi^{(2)}$	1	-1	-1	-1	1	1
$\Phi^{(3)}$	$\begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix}$	$\begin{Vmatrix} 0 & 1 \\ 1 & 0 \end{Vmatrix}$	$\begin{Vmatrix} 0 & \varepsilon \\ \varepsilon^{-1} & 0 \end{Vmatrix}$	$\begin{Vmatrix} 0 & \varepsilon^{-1} \\ \varepsilon & 0 \end{Vmatrix}$	$\begin{Vmatrix} \varepsilon & 0 \\ 0 & \varepsilon^{-1} \end{Vmatrix}$	$\begin{Vmatrix} \varepsilon^{-1} & 0 \\ 0 & \varepsilon \end{Vmatrix}$

EXEMPLE 2.— Une représentation orthogonale naturelle d'un groupe infini, à savoir de SU (2), est fournie par l'épimorphisme Φ : SU (2) $\mapsto$ SO (3) que nous avons construit encore au chapitre 7, § 1.

2. Réductibilité complète.— Les définitions et les remarques faites au § 1 montrent clairement le caractère fondamental de l'assertion suivante:

THÉORÈME 2 (théorème de Maschke). — *Toute représentation linéaire d'un groupe fini G sur un corps commutatif K de caractéristique un nombre ne divisant pas $|G|$ (en particulier de caractéristique nulle) est complètement réductible.*

Rappelons que l'assertion du théorème 2 exprime que (Φ, V) est décomposable en somme directe des représentations irréductibles. A proprement parler, le théorème classique de Maschke s'énonce comme suit :

(M). *Tout sous-espace G-invariant $U \subset V$ possède un supplémentaire G-invariant W :*

$$V = U \oplus W. \tag{3}$$

Nous allons démontrer justement cette dernière assertion dont le théorème 2 découle automatiquement. En effet, ou bien la représentation (Φ, V) est irréductible et alors il n'y a rien à démontrer, ou bien il existe un sous-espace strict G-invariant U et alors la décomposition (3), avec un certain G-sous-espace W, est vraie. Dans ce dernier cas, $\dim U < \dim V$, $\dim W < \dim V$. En appliquant à U et W les mêmes raisonnements et en utilisant la récurrence sur la dimension, on obtient la décomposition cherchée en composantes irréductibles. ■

Passons maintenant à la démonstration de l'assertion (M). Puisque nous nous intéressons toujours au cas du corps $K = \mathbb{C}$, il est utile de développer deux raisonnements indépendants.

Première démonstration $(K = \mathbb{C})$. — D'après le théorème 1, il existe sur l'espace de représentation V une forme hermitienne non dégénérée $(u \mid v)$ invariante par rapport aux opérateurs linéaires $\Phi(g)$. Pour chaque sous-espace $U \subset V$, il existe un *supplémentaire orthogonal*

$$U^\perp = \{v \in V \mid (u \mid v) = 0, \quad \forall u \in U\},$$

et, d'après un théorème connu en Algèbre linéaire,

$$V = U \oplus U^\perp,$$

avec $(U^\perp)^\perp = U$. Supposons maintenant que U soit un G-sous espace de V, c'est-à-dire que $\Phi(g)\, U \subset U$ pour tout $g \in G$. Comme $\Phi(g)|_U$ est un automorphisme, tout élément $u \in U$ s'écrit sous la forme $u = \Phi(g)\, u'$, $u' \in U$. Il ne reste qu'à utiliser la propriété d'une forme invariante :

$$v \in U^\perp \Rightarrow (u \mid \Phi(g)\, v) = (\Phi(g)\, u' \mid \Phi(g)\, v) = (u' \mid v) = 0.$$

Par conséquent, $v \in U^\perp \Rightarrow \Phi(g)\, v \in U^\perp$. En posant $W = U^\perp$, on obtient la décomposition (3). ■

Deuxième démonstration. — Soit, comme précédemment, U un sous-espace de V, invariant par l'opération de G. Considérons la somme directe

$$V = U \oplus U',$$

où U' est un supplémentaire de U arbitrairement choisi. En général, U' n'est pas un sous-espace G-invariant. Considérons l'opérateur $\mathcal{P}: V \to U'$ appelé projecteur défini pour tout vecteur $v =$

$= u + u'$ par la relation
$$\mathcal{P}v = u'.$$
On a
$$v - \mathcal{P}v \in U, \quad \mathcal{P}(U) = 0, \quad \mathcal{P}^2 = \mathcal{P}. \tag{4}$$
Introduisons maintenant un opérateur linéaire « moyen »:
$$\mathcal{P}_G = |G|^{-1} \sum_{h \in G} \Phi(h) \mathcal{P} \Phi(h^{-1})$$
(par hypothèse, la division par $|G|$ est possible). On a
$$\Phi(g) \mathcal{P}_G = \mathcal{P}_G \Phi(g), \quad \forall g \in G. \tag{5}$$
En effet
$$\Phi(g) \mathcal{P}_G \Phi(g^{-1}) = |G|^{-1} \sum_{h \in G} \Phi(g) \Phi(h) \mathcal{P} \Phi(h^{-1}) \Phi(g^{-1}) =$$
$$= |G|^{-1} \sum_{h \in G} \Phi(gh) \mathcal{P} \Phi((gh)^{-1}) =$$
$$= |G|^{-1} \sum_{t \in G} \Phi(t) \mathcal{P} \Phi(t^{-1}) = \mathcal{P}_G,$$
ce qui conduit à la relation (5). Posons
$$W = \mathcal{P}_G(V) = \{\mathcal{P}_G v \mid v \in V\}.$$
D'après (5), on a $\Phi(g) w = \Phi(g) \mathcal{P}_G v = \mathcal{P}_G \Phi(g) v = \mathcal{P}_G v' = w' \in W$ pour tout $w \in W$, si bien que le sous-espace vectoriel $W \subset V$ est réellement un G-sous-espace.

Il reste à montrer que $V = U \oplus W$ est somme directe de G-sous-espaces. Puisque $\Phi(h^{-1}) v - \mathcal{P}\Phi(h^{-1}) v \in U$ (voir (4)), on a $v - \Phi(h) \mathcal{P} \Phi(h^{-1}) v = \Phi(h) \{\Phi(h^{-1}) v - \mathcal{P}\Phi(h^{-1}) v\} \in \Phi(h) U = U$ (U est invariant). Par suite,
$$v - \mathcal{P}_G v = |G|^{-1} \sum_{h \in G} (v - \Phi(h) \mathcal{P} \Phi(h^{-1})) = u \in U,$$
et l'on obtient $v = u + w$, avec $w = \mathcal{P}_G v \in W$, c'est-à-dire $V = U + W$.

On a ensuite $\Phi(h^{-1}) U \subset U \Rightarrow \mathcal{P}\Phi(h^{-1}) U = 0$ (voir (4)) $\Rightarrow \Phi(h) \mathcal{P}\Phi(h^{-1}) U = 0 \Rightarrow \mathcal{P}_G(U) = 0$. Par conséquent, $v - \mathcal{P}_G v = u \in U \Rightarrow \mathcal{P}_G(v - \mathcal{P}_G v) = 0$, d'où $\mathcal{P}_G v = \mathcal{P}_G^2 v$ pour tout $v \in V$. Cela signifie que $\mathcal{P}_G$ est l'opération de projection sur W suivant U:
$$\mathcal{P}_G(U) = 0, \quad \mathcal{P}_G^2 = \mathcal{P}_G. \tag{6}$$
Maintenant, $v \in U \cap W \Rightarrow \mathcal{P}_G v = 0$, car $v \in U$, et $v = \mathcal{P}_G v'$, puisque $v \in \mathcal{P}_G(V) = W$. En utilisant (6) on obtient $0 = \mathcal{P}_G v = \mathcal{P}_G(\mathcal{P}_G v') = \mathcal{P}_G^2 v' = \mathcal{P}_G v' = v \Rightarrow U \cap W = 0$. ∎

Il serait imprudent de faire une conclusion plus forte sur l'unicité de la décomposition en composantes irréductibles (en G-sous-espaces irréductibles) : $V = V_1 \oplus V_2 \oplus \ldots \oplus V_r$. Si, par exemple, $\Phi(g) = \mathcal{E}$ est l'opérateur identique pour tout $g \in G$, toute décom-

position de V en sous-espaces de dimension un sera une décomposition en composantes irréductibles, le nombre de telles décompositions étant infiniment grand. Il n'en est plus de même si nous groupons les composantes irréductibles isomorphes:

$$V = U_1 \oplus \ldots \oplus U_s.$$

Puisque nous ne distinguons pas entre les G-espaces isomorphes, on peut poser

$$U_1 = V_1 \oplus V_1 \oplus \ldots \oplus V_1 = n_1 V_1,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$U_s = V_s \oplus V_s \oplus \ldots \oplus V_s = n_s V_s,$$

où n_i est l'*ordre de multiplicité* avec lequel la composante irréductible V_i figure dans la décomposition de V. Nous verrons plus loin que ces multiplicités sont définies de façon unique.

EXERCICES

1. Toute représentation continue de dimension un du groupe $(\mathbb{R}, +)$ (lorsqu'à des nombres voisins correspondent des opérateurs voisins) est de la forme $\Phi^{(\alpha)}: t \mapsto e^{i\alpha t}$, où α est un nombre complexe. Montrer que $\Phi^{(\alpha)}$ est unitaire si, et seulement si, $\alpha \in \mathbb{R}$. (Indication. Dériver l'égalité $e^{i\alpha t} \overline{e^{i\alpha t}} = 1$ par rapport à t et poser $t = 0$.)

2. Le noyau de l'homomorphisme $f: t \mapsto \begin{Vmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{Vmatrix}$ du groupe $(\mathbb{R}, +)$ sur SO (2) se compose des nombres $t = 2\pi m$, $m \in \mathbb{Z}$. Ainsi, SO (2) $\cong \mathbb{R}/2\pi\mathbb{Z}$, et à toute représentation unitaire irréductible Φ du groupe SO (2) (d'après les résultats du § 4, elle est nécessairement de dimension un) correspond une représentation unitaire irréductible $\tilde{\Phi}$: $t + 2\pi m \mapsto \Phi(t)$, $0 \leqslant t < 2\pi$, du groupe $\mathbb{R}$ pour laquelle $\tilde{\Phi}(2\pi) = \Phi(0) = 1$. Déduire de l'exercice 1 que $\tilde{\Phi} = \Phi^{(n)}$, $n \in \mathbb{Z}$. Jointe à la remarque 3) du nº 1, cette égalité signifie que toute représentation irréductible du groupe SO (2) est de la forme $\Phi^{(n)}(t) = e^{int}$, $n \in \mathbb{Z}$. Vérifier que

$$\frac{1}{2\pi}\int_0^{2\pi} e^{ikt} \cdot \overline{e^{ilt}}\, dt = \delta_{kl}$$

(comparer avec la relation obtenue dans l'exercice 3 du § 1: l'ordre n est remplacé par le « volume » 2π du groupe SO (2)). En Analyse, le système de fonctions $\{e^{int}\}$ fournit un exemple classique de système orthonormé complet de fonctions périodiques (ou de fonctions sur la circonférence $S^1 \sim$ SO(2)). C'est le point de départ pour la théorie bien développée des séries de Fourier.

3. En partant du théorème de Maschke, démontrer que toute représentation complexe exacte de dimension deux d'un groupe fini non abélien est irréductible.

§ 3. Groupes finis des rotations

Ce paragraphe est consacré aux sous-groupes finis du groupe SO (3). Leur connaissance nous permettra d'obtenir en même temps les représentations orthogonales irréductibles des groupes tels que

A_4, S_4, A_5 et cela sous une enveloppe géométrique facile à retenir. Lors d'une première lecture on peut omettre le n° 1 et la démonstration (assez schématique) du théorème 2, mais à tous ceux qui désirent s'assurer qu'ils ont bien assimilé l'idée générale des « opérations des groupes » (chap. 7, § 2) il sera utile d'étudier le contenu de tout le paragraphe.

1. Ordres des sous-groupes finis de SO (3).— D'après le théorème d'Euler étudié dans le cours d'Algèbre linéaire, tout élément $\mathcal{A} \in$ $\in$ SO (3), $\mathcal{A} \neq \mathcal{E}$, représente une rotation dans l'espace euclidien $\mathbb{R}^3$ autour d'un certain axe. Autrement dit, il existe exactement deux points sur la sphère unité S^2 de dimension deux, qui restent fixes lors de l'opération $\mathcal{A}$: ce sont les points d'intersection de la sphère et de l'axe de rotation. Ces deux points sont appelés *pôles de la rotation* $\mathcal{A}$.

Soient maintenant G un sous-groupe fini de SO (3) et S l'ensemble des pôles de toutes les rotations $\mathcal{A} \neq \mathcal{E}$ de G. Il est clair que G opère comme le groupe de permutations sur l'ensemble S. Si x est le pôle d'une certaine rotation $\mathcal{A} \neq \mathcal{E}$, $\mathcal{A} \in G$, on a pour tout $\mathcal{B} \in G$:

$$(\mathcal{B}\mathcal{A}\mathcal{B}^{-1})\, \mathcal{B}x = \mathcal{B} \cdot \mathcal{A}x = \mathcal{B}x,$$

c'est-à-dire que $\mathcal{B}x$ est le pôle de $\mathcal{B}\mathcal{A}\mathcal{B}^{-1}$, et donc $\mathcal{B}x \in S$. Désignons par Ω l'ensemble de tous les couples $(\mathcal{A}, x)$, où $\mathcal{A} \in G$, $\mathcal{A} \neq \mathcal{E}$ et x est un pôle de $\mathcal{A}$. Soit ensuite G_x le stabilisateur (sous-groupe stationnaire) du point x, c'est-à-dire le sous-groupe de G comprenant tous les éléments qui laissent fixe le point x. Si

$$G = G_x \cup g_2 G_x \cup \ldots \cup g_{m_x} G_x$$

est la décomposition de G en classes à gauche suivant G_x, alors la G-orbite du point x sera l'ensemble

$$G(x) = \{x, g_2 x, \ldots, g_{m_x} x\}$$

dont le nombre d'éléments est $|G(x)| = m_x$. Suivant le théorème de Lagrange, $N = m_x n_x$, où $N = |G|$, $n_x = |G_x|$ (par rapport au chap. 7, § 1, les désignations sont légèrement modifiées). Remarquons que n_x est l'ordre d'un sous-groupe cyclique de G dont chacun des éléments est une rotation autour de l'axe passant par x. On dit que n_x est l'*ordre de multiplicité du pôle* x ou que x est un n_x-*pôle*.

Vu qu'à tout élément $\mathcal{A} \neq \mathcal{E}$ de G correspondent deux pôles, on a $|\Omega| = 2(N - 1)$.

D'autre part, pour chaque pôle x, il existe $n_x - 1$ éléments de G différant de e et laissant fixe le pôle x. Par conséquent, le nombre

des couples $(\mathscr{A}, x)$ est égal à la somme

$$|\Omega| = \sum_{x \in S} (n_x - 1).$$

Prenons pour $\{x_1, \ldots, x_k\}$ l'ensemble des pôles, un pôle de chaque orbite. En posant $n_i = n_{x_i}$, $m_i = m_{x_i}$ et en remarquant que $n_x = n_{x_i} = n_i$ pour tout $x \in G(x_i)$, nous obtenons

$$|\Omega| = \sum_{x \in S} (n_x - 1) = \sum_{i=1}^{k} m_i (n_i - 1) = \sum_{i=1}^{k} (N - m_i).$$

Ainsi,

$$2N - 2 = \sum_{i=1}^{k} (N - m_i).$$

En divisant par N les deux membres de l'égalité, nous aurons

$$2 - \frac{2}{N} = \sum_{i=1}^{k} \left(1 - \frac{1}{n_i}\right). \tag{1}$$

Supposons $N > 1$, de sorte que $1 \leqslant 2 - \frac{2}{N} < 2$. Puisque $n_i \geqslant 2$, on a $\frac{1}{2} \leqslant 1 - \frac{1}{n_i} < 1$ et, de ce fait, k doit être égal à 2 ou à 3.

Cas 1. $k = 2$. Alors

$$2 - \frac{2}{N} = \left(1 - \frac{1}{n_1}\right) + \left(1 - \frac{1}{n_2}\right),$$

ou ce qui revient au même

$$2 = \frac{N}{n_1} + \frac{N}{n_2} = m_1 + m_2,$$

d'où $m_1 = m_2 = 1$, $n_1 = n_2 = N$. Par suite, G possède un axe de rotation et un seul, et $G = C_N$ est un groupe cyclique d'ordre N.

Cas 2. $k = 3$. Soit, pour plus de détermination, $n_1 \leqslant n_2 \leqslant n_3$. Si $n_1 \geqslant 3$, nous aurions

$$\sum_{i=1}^{3} \left(1 - \frac{1}{n_i}\right) \geqslant \sum_{i=1}^{3} \left(1 - \frac{1}{3}\right) = 2,$$

ce qui est impossible. Ainsi, $n_1 = 2$ et l'équation (1) prend la forme

$$\frac{1}{2}+\frac{2}{N}=\frac{1}{n^2}+\frac{1}{n^3}.$$

Il est évident que $n_2 \geqslant 4 \Rightarrow \frac{1}{n_2}+\frac{1}{n_3} \leqslant \frac{1}{2}$ est une contradiction. Par conséquent $n_2 = 2$ ou 3.

Si $n_2 = 2$, alors $n_3 = \frac{N}{2} = m$ (N doit être pair) et $m_1 = m_2 = m$, $m_3 = 2$. Ces données correspondent au groupe diédral D_m (voir chap. 7, § 3, nº 5, exemple 1).

Si $n_2 = 3$, alors

$$\frac{1}{6}+\frac{2}{N}=\frac{1}{n_3},$$

et nous n'avons que trois possibilités:

2′) $n_3 = 3$, $N = 12$, $m_1 = 6$, $m_2 = 4$, $m_3 = 4$;
2″) $n_3 = 4$, $N = 24$, $m_1 = 12$, $m_2 = 8$, $m_3 = 6$;
2‴) $n_3 = 5$, $N = 60$, $m_1 = 30$, $m_2 = 20$, $m_3 = 12$.

Groupons toutes ces données sur la table ci-dessous:

N	Nombres d'orbites	$\vert S \vert$	Ordres des stabilisateurs		
n	2	2	n	n	—
$2m$	3	$2m+2$	2	2	m
12	3	14	2	3	3
24	3	26	2	3	4
60	3	62	2	3	5

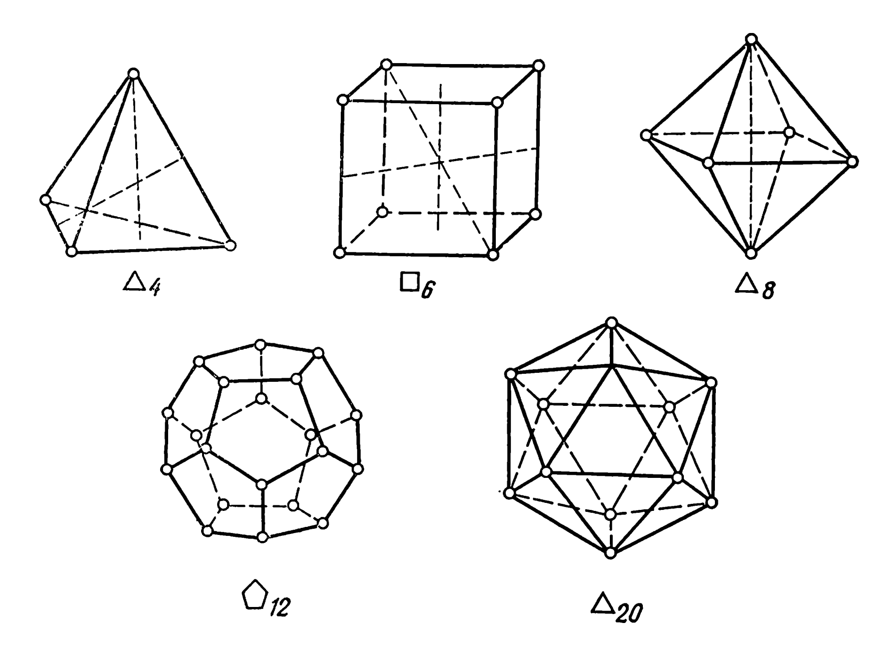
(2)

Ainsi, nous avons démontré l'assertion suivante:

THÉORÈME 1. — *Soit G un sous-groupe fini de* SO (3), *distinct d'un groupe cyclique et d'un groupe diédral. Alors, il n'existe pour son ordre que trois possibilités*: $N = 12$, 24 *ou* 60. *D'autres limitations imposées au groupe G sont contenues dans la table* (2). ∎

2. Groupes des polyèdres réguliers.— L'existence des groupes d'ordre 12, 24 et 60 contenus dans SO (3) et distincts des groupes cycliques et diédraux se démontre de façon bien simple. Dans l'espace euclidien $\mathbb{R}^3$, il n'existe, à une similitude près, que cinq polyèdres convexes réguliers (connus depuis l'Antiquité): le tétraèdre

$\triangle_4$, le cube $\square_6$, l'octaèdre $\triangle_8$, le dodécaèdre $⬠_{12}$ et l'icosaèdre $\triangle_{20}$:



Si le centre d'un polyèdre régulier M est placé à l'origine de l'espace $\mathbb{R}^3$, les rotations de SO (3), qui amènent M en coïncidence avec lui-même, formeront un sous-groupe fini. Pourtant, au lieu de cinq il n'y a dans ce cas que trois groupes des rotations différents (non isomorphes), vu qu'ils sont identiques pour le cube et l'octaèdre ainsi que pour le dodécaèdre et l'icosaèdre. Géométriquement, cela s'explique très simplement. Si l'on joint par des segments les milieux des faces adjacentes d'un cube, ces segments seront des arêtes d'un octaèdre inscrit dans le cube. Toute rotation dans $\mathbb{R}^3$, qui laisse invariant le cube, transforme en lui-même l'octaèdre inscrit, et réciproquement. Une remarque analogue est aussi applicable au couple dodécaèdre-icosaèdre. Dans la table ci-dessous, N_0 est le nombre de sommets d'un polyèdre, N_1 le nombre d'arêtes, N_2 le nombre de faces, μ le nombre de côtés (d'arêtes) de chaque face et ν le nombre de faces aboutissant à un même sommet. Comme précédemment, N est l'ordre du groupe correspondant.

Suivant le théorème géométrique d'Euler sur les polyèdres, $N_0 - N_1 + N_2 = 2$. Le nombre total de pôles est égal à $N_0 + N_1 + N_2 = 2N_1 + 2$. Lors de toute rotation qui transforme un polyèdre en lui-même, une arête donnée a_1b_1 vient en coïncidence avec toute autre arête a_ib_i ou b_ia_i, de sorte que $N = 2N_1$. Remar-

quons encore que $\{\mu, \nu\} = \{n_2, n_3\}$, où n_2, n_3 sont les ordres de multiplicité des pôles que nous avons introduits au nº 1.

Soient ensuite **T** le groupe du tétraèdre, **O** le groupe du cube (de l'octaèdre) et **I** le groupe de l'icosaèdre (du dodécaèdre).

	N_0	N_1	N_2	μ	ν	N
Tétraèdre . . . . $\triangle_4$	4	6	4	3	3	12
Cube . . . . $\square_6$	8	12	6	4	3	24
Octaèdre . . . . $\triangle_8$	6	12	8	3	4	24
Dodécaèdre . . . . ⬠$_{12}$	20	30	12	5	3	60
Icosaèdre . . . . $\triangle_{20}$	12	30	20	3	5	60

Les éléments de **T** sont les rotations de certains angles autour de quatre axes qui relient les sommets aux centres des faces opposées, ainsi que les rotations de l'angle π autour de chacun des trois axes qui passent par les milieux des arêtes opposées, et la rotation unité.

Le groupe **O** contient, en plus de la rotation unité, les rotations d'angles $\pi/2$, π, $3\pi/2$ autour de trois axes qui joignent les centres des faces opposées du cube, les rotations d'angles $2\pi/3$, $4\pi/3$ autour de quatre axes qui relient les sommets opposés n'appartenant pas à la même face, et les rotations d'angle π autour de chacun de six axes qui relient les milieux des arêtes diagonalement opposées.

Un tétraèdre régulier s'inscrit dans un cube et reste invariant pour certaines rotations de **O** d'ordre 2 ou 3. Ces rotations dont le nombre, avec la rotation unité, est égal à 12, constituent justement le groupe **T**. Par conséquent, $\mathbf{T} \subset \mathbf{O}$, et puisque $|\mathbf{O} : \mathbf{T}| = 2$, on a $\mathbf{T} \lhd \mathbf{O}$.

A chaque élément de **O** correspond une permutation et une seule sur l'ensemble constitué de quatre grandes diagonales du cube. L'égalité des ordres des groupes $|\mathbf{O}| = |S_4| = 24$ entraîne leur isomorphisme: $\mathbf{O} \cong S_4$. Respectivement, $\mathbf{T} \cong A_4$. L'exercice 2 montre que $\mathbf{I} \cong A_5$.

En revenant à la démonstration du théorème 1, remarquons que pour $n_1 = 2$, $n_2 = n_3 = 3$, il existe deux orbites à quatre éléments $G(p_1) = \{p_1, p_2, p_3, p_4\}$, $G(q_1) = \{q_1, q_2, q_3, q_4\}$ des pôles, où p_i et q_i sont des points opposés sur la sphère S^2. Si $\triangle_4^\bullet$ est un tétraèdre à sommets p_i, son groupe de transformations de symétrie $\mathbf{T}^\circ$ contient G. De $|G| = 12$ il ressort que $\triangle_4^\bullet$ est un tétraèdre régulier, c'est-à-dire que $\triangle_4^\circ = \triangle_4$ et $\mathbf{T}^\circ = G = \mathbf{T}$.

Pour $n_2 = 3$, $n_3 = 4$, prenons l'orbite à six éléments $G(p_1) = \{p_1, \ldots, p_6\}$ des pôles qui se répartissent en couples, car $i \neq 3 \Rightarrow n_i \neq 4$. Ces trois couples de points sur la sphère S^2 seront

pris pour trois couples de sommets opposés d'un octaèdre $\triangle_8^\circ$. De même que dans le cas précédent, $|G| = 24 \Rightarrow \triangle_8^\circ = \triangle_8$ (en ce sens que $\triangle_8^\circ$ est un octaèdre régulier) et $\mathbf{O}^\circ = G = \mathbf{O}$.

Enfin pour $n_1 = 2$, $n_2 = 3$, $n_3 = 5$, on construit un icosaèdre $\triangle_{20}^\circ$ de sommets p_i pris sur l'orbite $G(p_1) = \{p_1, \ldots, p_{20}\}$. De nouveau, $|G| = 60$ implique que $\triangle_{20}^\circ$ est un icosaèdre régulier et que les groupes coïncident: $\mathbf{I}^\circ = G = \mathbf{I}$.

Il reste à remarquer que quels que soient deux polyèdres réguliers de même type inscrits dans la sphère S^2, ils sont obtenus l'un à partir de l'autre par une certaine rotation (un changement de système de coordonnées). C'est ainsi que dans SO (3) les sous-groupes isomorphes deviennent conjugués. Résumons les résultats obtenus sous forme d'un théorème.

THÉORÈME 2. — *Les seuls sous-groupes finis de* SO (3) *sont, à un isomorphisme près, les groupes* C_n, D_n, $n \in \mathbb{N}$; $\mathbf{T} \cong A_4$, $\mathbf{O} \cong S_4$ *et* $\mathbf{I} \cong A_5$. *Deux sous-groupes finis isomorphes quelconques sont conjugués dans* SO (3). ■

COROLLAIRE. — *Les isomorphismes indiqués dans le théorème* 2 *donnent des représentations orthogonales irréductibles de dimension trois des groupes* A_4, S_4 *et* A_5. ■

En utilisant le théorème 2 et l'épimorphisme Φ: SU (2) → SO (3) (chap. 7, § 1, théorème 1), on arrive facilement à la description de tous les sous-groupes finis du groupe SU (2) (on peut aussi opérer dans l'ordre inverse). Un tel groupe G^*, différent d'un groupe cyclique, est l'image réciproque d'un certain sous-groupe fini $G \subset$ SO (3), C'est ainsi qu'apparaissent les *groupes* dits *binaires*:

$$D_n^* = \Phi^{-1}(D_n), \quad \mathbf{T}^* = \Phi^{-1}(\mathbf{T}), \quad \mathbf{O}^* = \Phi^{-1}(\mathbf{O}), \quad \mathbf{I}^* = \Phi^{-1}(\mathbf{I}),$$

le groupe diédral binaire, le groupe binaire du tétraèdre, le groupe binaire de l'octaèdre et le groupe binaire de l'icosaèdre. Les groupes binaires, de même que les représentations orthogonales Φ: SU (2) → SO (3), apparaissent, dans leur ensemble, de façon naturelle lors de la description des états d'un système physique de particules douées de spin.

EXERCICES

1.. En plus du sous-groupe unité, le groupe de l'icosaèdre **I** comprend 15 sous-groupes cycliques conjugués d'ordre 2; 10 sous-groupes cycliques conjugués d'ordre 3 et 6 sous-groupes cycliques conjugués d'ordre 5. Démontrer que **I** est un groupe simple (I n d i c a t i o n. Relire la démonstration du théorème 5 du chap. 7, § 3.)

2. Etablir un isomorphisme entre les groupes **I** et A_5. (I n d i c a t i o n. En utilisant le fait que tous les éléments d'ordre 2 sont conjugués, montrer qu'ils se disposent dans un « bouquet » (fig. 21) de cinq sous-groupes de Sylow conjugués, disjoints deux à deux (ou plus exactement présentant une intersec-

tion suivant e) d'ordre 4. Le groupe **I** opère sur le « bouquet » par conjugaison. Cette opération est exacte, car **I** est un groupe simple (voir exercice 1).)

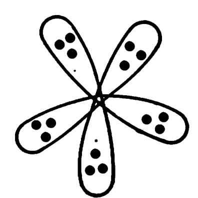


Fig. 21.

3. Si H est un sous-groupe fini d'ordre impair de SU (2) ou de SO (3), il est cyclique. (Indication. Appliquer le théorème d'homomorphie au cas $\Phi: SU(2) \to SO(3)$.)

4. Montrer que, si un sous-groupe fini $H \subset$ SU (2) n'est pas une image réciproque d'un sous-groupe quelconque $G \subset$ SO (3), alors $|H| \equiv 1 \pmod{2}$.

5. Montrer qu'à une conjugaison près

$$D_3^* = \left\langle \begin{Vmatrix} 0 & 1 \\ -1 & 0 \end{Vmatrix}, \begin{Vmatrix} \varepsilon & 0 \\ 0 & \varepsilon^{-1} \end{Vmatrix} \middle| \varepsilon^2+\varepsilon+1=0 \right\rangle,$$

6. Qu'y a-t-il de commun entre le groupe binaire de l'icosaèdre $\mathbf{I}^*$ et le groupe

$$\mathrm{SL}(2, Z_5) = \left\{ \begin{Vmatrix} a & b \\ c & d \end{Vmatrix} \middle| ad-bc=1;\ a, b, c, d \in Z_5 \right\} ?$$

7. Soient des atomes de q espèces différentes ($q < 200$), qui se disposent de toutes les façons possibles aux sommets d'un polyèdre régulier M (on ne tient compte d'aucune liaison chimique). Les « molécules » obtenues l'une à partir de l'autre par rotation autour d'un certain axe ne sont pas discernées. Soit $f(M, q)$ le nombre de « molécules » différentes. Obtenir les formules

$$f(\triangle_4, q) = \frac{q^2}{12}(q^2+11),$$

$$f(\square_6, q) = \frac{q^2}{24}(q^6+17q^2+6),$$

$$f(\triangle_8, q) = \frac{q^2}{24}(q^4+3q^2+12q+8).$$

(Indication. Appliquer les considérations utilisées lors du calcul du nombre de colliers (problème 2 au début du chapitre).)

8. Montrer que le calcul du nombre de différentes peintures des faces du polyèdre M avec q couleurs conduit dans le cas d'un tétraèdre $\triangle_4$ à la même formule que dans l'exercice 7, alors que dans le cas d'un cube et d'un octaèdre les formules changent de place.

§ 4. Caractères des représentations linéaires

1. Lemme de Schur et son corollaire.— La base de toute théorie mathématique cohérente est généralement constituée par quelques considérations relativement peu compliquées (mais bien fines). L'une des pierres angulaires de la théorie des représentations est l'assertion suivante:

THÉORÈME 1 (lemme de Schur). — *Soient* (Φ, V), (Ψ, W) *deux représentations complexes irréductibles d'un groupe* G, *et* $\sigma: V \to W$ *une application linéaire telle que*

$$\Psi(g)\sigma = \sigma\Phi(g), \quad \forall g \in G. \tag{1}$$

Alors:

(i) $\sigma = 0$ *si les représentations* Φ, Ψ *ne sont pas équivalentes*;
(ii) $\sigma = \lambda\mathscr{E}$ *si* $V = W$ *et* $\Phi = \Psi$.

DÉMONSTRATION. — Si $\sigma = 0$, il n'y a rien à démontrer. Considérons donc $\sigma \neq 0$ et posons $V_0 = \operatorname{Ker} \sigma \subset V$.

Puisque $\sigma\Phi(g) v_0 = \Psi(g) \sigma v_0 = 0$ pour tout $v_0 \in V_0$, on a $\Phi(g) V_0 = V_0$, c'est-à-dire le sous-espace V_0 est invariant par G. Du fait que (Φ, V) est irréductible, on déduit que $V_0 = 0$ ou $V_0 = V$. L'égalité $V_0 = V$ est impossible, car $\sigma \neq 0$. Par suite, $\operatorname{Ker} \sigma = 0$.

D'une manière analogue, en posant $W_1 = \operatorname{Im} \sigma \subset W$, on aura $w_1 \in W_1 \Rightarrow \Psi(g) w_1 = \Psi(g)\sigma(v_1) = \sigma\ (\Phi(g) v_1) = w_1' \in W_1$, et donc W_1 est un sous-espace invariant de W. De nouveau, $\sigma \neq 0 \Rightarrow W_1 \neq 0$, et puisque (Ψ, W) est une représentation irréductible, il ne reste qu'une seule possibilité $W_1 = W$.

(i) Puisque $\operatorname{Ker} \sigma = 0$, $\operatorname{Im} \sigma = W$, alors $\sigma: V \to W$ est un isomorphisme, et la condition (1) n'est rien d'autre que la condition d'équivalence des représentations Φ, Ψ (voir définition 2 au § 1). L'assertion (i) est démontrée.

(ii) Par hypothèse, $\sigma: V \to V$ est un opérateur linéaire sur V. Soit λ l'une de ses valeurs propres; elle existe parce que le corps de base $\mathbb{C}$ est algébriquement clos. L'opérateur linéaire $\sigma_0 = \sigma - \lambda\mathscr{E}$ possède un noyau non trivial (il contient le vecteur propre) et satisfait à l'égalité $\Psi(g)\sigma_0 = \sigma_0\Phi(g)$. Du fait de ce qui a été démontré, $\sigma_0 = 0$, c'est-à-dire $\sigma = \lambda\mathscr{E}$. ∎

COROLLAIRE. — *Soient* (Φ, V), (Ψ, W) *deux représentations irréductibles sur* $\mathbb{C}$ *d'un groupe fini* G *d'ordre* $|G|$, *et* $\sigma: V \to W$ *une application linéaire arbitraire. Alors, l'application « moyenne »*

$$\tilde{\sigma} = \frac{1}{|G|} \sum_{g \in G} \Psi(g)\sigma\Phi(g)^{-1}$$

possède les propriétés suivantes:

(i) $\Phi \not\approx \Psi \Rightarrow \tilde{\sigma} = 0$;

(ii) $V = W,\ \Phi = \Psi \Rightarrow \tilde{\sigma} = \lambda\mathscr{E}, \quad \lambda = \dfrac{\operatorname{tr} \sigma}{\dim V}$.

DÉMONSTRATION. — On a

$$\Psi(g)\tilde{\sigma}\Phi(g)^{-1} = |G|^{-1} \sum_{h \in G} \Psi(g)\Psi(h)\sigma\Phi(h)^{-1}\Phi(g)^{-1} =$$

$$= |G|^{-1} \sum_{n} \Psi(gh)\sigma\Phi(gh)^{-1} = |G|^{-1} \sum_{t \in G} \Psi(t)\sigma\Phi(t)^{-1} = \tilde{\sigma}$$

de sorte que $\Psi(g)\tilde{\sigma} = \tilde{\sigma}\Phi(g)$, $\forall g \in G$. D'après le lemme de Schur, on obtient tout de suite les deux assertions, la précision relative à la constante λ découlant des relations

$$(\dim V)\lambda = \operatorname{tr}\lambda\mathscr{E} = \operatorname{tr}\tilde{\sigma} = |G|^{-1}\sum_{g\in G}\operatorname{tr}\Phi(g)\sigma\Phi(g)^{-1} =$$

$$= |G|^{-1}\sum_{g\in G}\operatorname{tr}\sigma = \operatorname{tr}\sigma.$$

Nous avons utilisé ici la propriété connue de la fonction de trace: $\operatorname{tr} CAC^{-1} = \operatorname{tr} A$. ■

Nous aurons besoin de la *traduction matricielle du corollaire*. A cet effet, choisissons dans les espaces V, W des bases quelconques: $V = \langle e_i \mid i \in I\rangle$, $W = \langle f_j \mid j \in J\rangle$. Ecrivons dans ces bases nos applications linéaires (en les identifiant avec les matrices correspondantes):

$$\Phi_g = (\varphi_{ii'}(g)), \quad \Psi_g = (\psi_{jj'}(g)),$$
$$\sigma = (\sigma_{ji}), \quad \tilde{\sigma} = (\tilde{\sigma}_{ji});\ i,\ i' \in I,\ j,\ j' \in J.$$

Par la définition même de $\tilde{\sigma}$, on a

$$\tilde{\sigma}_{ji} = |G|^{-1}\sum_{g\in G}\sum_{i'\in I,\ j'\in J}\psi_{jj'}(g)\,\sigma_{j'i'}\varphi_{i'i}(g^{-1}). \tag{2}$$

L'application linéaire $\sigma: V \to W$ est tout à fait arbitraire. Nous pouvons prendre

$$\sigma_{ji} = 0,\quad \forall (j,\ i) \neq (j_0,\ i_0);\quad \sigma_{j_0 i_0} = 1. \tag{3}$$

Alors, à l'assertion (i) du corollaire correspond la relation

$$|G|^{-1}\sum_{g\in G}\psi_{jj_0}(g)\cdot\varphi_{i_0 i}(g^{-1}) = 0,\ \forall i,\ i_0,\ j,\ j_0 \tag{4}$$

(les représentations Φ et Ψ ne sont pas équivalentes).

Si maintenant $V = W$ et $\Phi = \Psi$, on a

$$\operatorname{tr}\sigma = \sum \sigma_{ii} = \sum_{i',\ j'}\delta_{j'i'}\sigma_{j'i'},$$

$$\tilde{\sigma} = \frac{\operatorname{tr}\sigma}{\dim V}\mathscr{E} \Rightarrow \tilde{\sigma}_{ji} = \delta_{ji}\frac{\operatorname{tr}\sigma}{\dim V} = \frac{\delta_{ji}}{\dim V}\sum_{i',\ j'}\delta_{j'i'}\sigma_{j'i'}.$$

En rapprochant la relation obtenue de la relation (2), on obtient

$$|G|^{-1}\sum_{g\in G,\ i',\ j'}\varphi_{jj'}(g)\,\sigma_{j'i'}\varphi_{i'i}(g^{-1}) = \frac{1}{\dim V}\sum_{i',\ j'}\delta_{ji}\delta_{j'i'}\sigma_{j'i'},$$

d'où, du fait que le choix de σ est arbitraire (voir (3)), nous concluons qu'à l'assertion (ii) du corollaire correspond la relation

$$|G|^{-1}\sum_{g\in G}\varphi_{jj_0}(g)\,\varphi_{i_0i}(g^{-1})=\begin{cases}\dfrac{\delta_{ji}}{\dim V} & \text{si } j_0=i_0,\\ 0 & \text{sinon.}\end{cases} \tag{5}$$

Les relations (4) et (5) contiennent toutes les informations dont nous avons besoin. ■

2. **Caractères des représentations.**— A chaque représentation complexe linéaire de dimension finie (Φ, V) d'un groupe G est associée une fonction

$$\chi_\Phi : G\to\mathbb{C},$$

définie par la relation

$$\chi_\Phi(g)=\operatorname{tr}\Phi(g),\quad g\in G,$$

et appelée *caractère de la représentation*. On la désigne aussi par le symbole χ_V ou tout simplement par χ s'il est clair de quelle représentation il s'agit.

Soient $\Phi_g=(\varphi_{ij}(g))$ une matrice correspondant à l'opérateur $\Phi(g)$ dans une certaine base de l'espace V et $\lambda_1, \ldots, \lambda_n$ ($n = \dim V$) ses racines caractéristiques comptées avec leur multiplicité.

Par définition,

$$\chi_\Phi(g)=\chi_V(g)=\sum_{i=1}^{n}\varphi_{ii}(g)=\sum_{i=1}^{n}\lambda_i.$$

Si C est une matrice inversible quelconque, on a

$$\operatorname{tr}C\Phi_gC^{-1}=\operatorname{tr}\Phi_g.$$

Or, nous savons que toute représentation Ψ équivalente à Φ est de la forme $g\mapsto C\Phi_gC^{-1}$. Par conséquent, les *caractères des représentations isomorphes* (*équivalentes*) *coïncident*. Cette remarque montre que la notion de caractère a été définie correctement.

Indiquons encore quelques propriétés élémentaires des caractères.

PROPOSITION. — *Soit χ_Φ le caractère d'une représentation complexe linéaire (Φ, V) d'un groupe G. Alors*:

(i) $\chi_\Phi(e)=\dim V$;

(ii) $\chi_\Phi(hgh^{-1})=\chi_\Phi(g)$, $\forall g, h\in G$, *c'est-à-dire* χ_Φ *est une fonction constante sur les classes des éléments conjugués du groupe* G;

(iii) $\chi_\Phi(g^{-1})=\overline{\chi_\Phi(g)}$ *pour tout élément* $g\in G$ *d'ordre fini* (*la barre signifie la conjugaison complexe*);

(iv) *à la somme directe* $\Phi=\Phi'\dotplus\Phi''$ *des représentations correspond le caractère* $\chi_\Phi=\chi_{\Phi'}+\chi_{\Phi''}$.

DÉMONSTRATION. — En effet, $\chi_\Phi(e) = \operatorname{tr}\Phi(e) = \operatorname{tr}\mathscr{E} = \dim V$. On a ensuite $\chi_\Phi(hgh^{-1}) = \operatorname{tr}\Phi(hgh^{-1}) = \operatorname{tr}\Phi(h)\Phi(g)\Phi(h)^{-1} = \operatorname{tr}\Phi(g) = \chi_\Phi(g)$. Pour démontrer (iii) remarquons que

$$g^m = e \Rightarrow \Phi(g)^m = \mathscr{E}$$

et que $\lambda_1^k, \ldots, \lambda_n^k$ sont les racines caractéristiques de l'opérateur $\Phi(g)^k$ si $\lambda_1, \ldots, \lambda_n$ sont les racines caractéristiques de l'opérateur $\Phi(g)^k$. En particulier, $\lambda_i^m = 1$, $1 \leqslant i \leqslant n$, et donc $|\lambda_i| = 1$, $\overline{\lambda}_i = \lambda_i^{-1}$. De ce fait

$$\chi_\Phi(g^{-1}) = \operatorname{tr}\Phi(g^{-1}) = \operatorname{tr}\Phi(g)^{-1} = \sum_i \lambda_i^{-1} = \sum_i \overline{\lambda}_i = (\overline{\sum_i \lambda_i}) = \overline{\chi_\Phi(g)}.$$

Enfin, dans le cas où $\Phi = \Phi' \dotplus \Phi''$, nous savons qu'avec un choix convenable de la base dans l'espace de représentation V toutes les matrices Φ_g, $g \in G$, prendront la forme

$$\Phi_g = \begin{Vmatrix} \Phi'_g & 0 \\ 0 & \Phi''_g \end{Vmatrix},$$

d'où $\operatorname{tr}\Phi_g = \operatorname{tr}\Phi'_g + \operatorname{tr}\Phi''_g$. Or, cela signifie justement que $\chi_\Phi(g) = \chi_{\Phi'}(g) + \chi_{\Phi''}(g)$. ■

Remarquons que pour $n = \dim V = 1$ on aura $\chi_\Phi(g) = \Phi(g)$, alors que pour $n > 1$ le caractère χ_Φ n'est pas un homomorphisme de G dans $\mathbb{C}^*$.

EXEMPLE 1.—Considérons le groupe SU (2) dans sa représentation naturelle de dimension deux. Soit χ le caractère correspondant. D'après la relation (5) du chapitre 7, § 1, toute matrice $g \in$ SU (2) est conjuguée à la matrice

$$b_\varphi = \begin{Vmatrix} e^{i\frac{\varphi}{2}} & 0 \\ 0 & e^{-i\frac{\varphi}{2}} \end{Vmatrix} \qquad 0 \leqslant \varphi < 2\pi,$$

de sorte que les classes des éléments conjugués du groupe SU (2) sont paramétrisées par des nombres réels φ pris dans l'intervalle indiqué. Conformément à la propriété (ii) des caractères on a

$$\chi(g) = \chi(Ub_\varphi U^{-1}) = \chi(b_\varphi) = e^{i\frac{\varphi}{2}} + e^{-i\frac{\varphi}{2}} = 2\cos\frac{\varphi}{2}.$$

Lors de la représentation canonique Φ: SU (2) → SO (3), la matrice b_φ se transforme en la matrice

$$B_\varphi = \begin{Vmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{Vmatrix},$$

qui constitue, elle aussi, un représentant commode de la classe des matrices orthogonales conjuguées du groupe SO (3). Il est évident que

$$\chi_\Phi (B_\varphi) = 1 + 2 \cos \varphi. \tag{6}$$

La formule (6) sera utilisée plus tard.

L'ensemble $\mathbb{C}^G = \{G \to \mathbb{C}\}$ de toutes les fonctions de G dans $\mathbb{C}$ est muni d'une structure naturelle d'espace vectoriel sur $\mathbb{C}$: pour $\alpha_1, \alpha_2 \in \mathbb{C}$, $\chi_1, \chi_2 \in \mathbb{C}^G$, on entend par $\alpha_1\chi_1 + \alpha_2\chi_2$ une fonction à valeurs

$$(\alpha_1\chi_1 + \alpha_2\chi_2)(g) = \alpha_1\chi_1(g) + \alpha_2\chi_2(g).$$

Une fonction de $\mathbb{C}^G$ est dite *centrale* si elle est constante sur les classes de conjugaison du groupe G. Les fonctions centrales forgment évidemment un sous-espace vectoriel de $\mathbb{C}^G$ que nous désignerons par le symbole $X_\mathbb{C}(G)$. En général, $X_\mathbb{C}(G)$ est un espace de dimension infinie, mais si le groupe G ne contient qu'un nombre fini de classes de conjugaison $C_1, C_2, \ldots, C_r$ (il ne sera toujours ainsi pour un groupe fini G), l'espace $X_\mathbb{C}(G)$ est de dimension finie. Par exemple,

$$X_\mathbb{C}(G) = \langle \Gamma_1, \Gamma_2, \ldots, \Gamma_r \rangle_\mathbb{C}, \tag{7}$$

où

$$\Gamma_i(g) = \begin{cases} 1 \text{ si } g \in C_i, \\ 0 \text{ si } g \notin C_i. \end{cases} \tag{7}$$

D'après ce qui a été démontré (la proposition (ii)), les caractères du groupe G appartiennent à l'espace $X_\mathbb{C}(G)$. Nous verrons plus loin que le sous-espace engendré par ces caractères coïncide en réalité avec $X_\mathbb{C}(G)$, en tout cas pour un groupe fini G.

Supposons plus loin que le groupe G soit fini. Transformons $\mathbb{C}^G$ en un espace hermitien muni du produit scalaire

$$(\sigma, \tau)_G = \frac{1}{|G|} \sum_{g \in G} \sigma(g)\overline{\tau(g)}, \quad \sigma, \tau \in \mathbb{C}^G. \tag{8}$$

On vérifie aisément que la forme $(\sigma, \tau) \mapsto (\sigma, \tau)_G$ possède toutes les propriétés d'une forme hermitienne non dégénérée. Sa restriction au sous-espace $X_\mathbb{C}(G) \subset \mathbb{C}^G$ constitue un instrument bien utile, surtout pour l'étude des caractères des représentations linéaires.

THÉORÈME 2. — *Soient Φ, Ψ deux représentations complexes irréductibles d'un groupe fini G. Alors,*

$$(\chi_\Phi, \chi_\Psi)_G = \begin{cases} 1 \text{ si } \Phi \approx \Psi, \\ 0 \text{ si } \Phi \not\approx \Psi. \end{cases} \tag{9}$$

DÉMONSTRATION. — En utilisant les notations matricielles, on a

$$\chi_\Phi(g) = \sum_{i=1}^{n} \varphi_{ii}(g), \quad \chi_\Psi(g) = \sum_{i=1}^{n} \psi_{ii}(g).$$

En posant $i_0 = i$, $j_0 = j$ dans la relation (4), puis, en effectuant la sommation sur i et j (dans des limites admissibles pour i et j), on obtient

$$\begin{aligned} 0 &= |G|^{-1} \sum_{g,\,i,\,j} \psi_{jj}(g)\,\varphi_{ii}(g^{-1}) = \\ &= |G|^{-1} \sum_{g} \Big(\sum_{j} \psi_{jj}(g)\Big)\Big(\sum_{i} \varphi_{ii}(g^{-1})\Big) = \\ &= |G|^{-1} \sum_{g\in G} \chi_\Psi(g)\,\chi_\Phi(g^{-1}) = \\ &= |G|^{-1} \sum_{g\in G} \chi_\Psi(g)\,\overline{\chi_\Phi(g)} = (\chi_\Psi, \chi_\Phi)_G, \end{aligned}$$

quelles que soient les représentations irréductibles non équivalentes Φ, Ψ du groupe G.

Utilisons maintenant (pour $i_0 = i$, $j_0 = j$) la relation (5):

$$\begin{aligned} 1 &= \Big(\sum_{j,\,i} \delta_{ji}\Big)/\dim V = |G|^{-1} \sum_{g\in G} \Big(\sum_{j} \varphi_{jj}(g)\Big)\Big(\sum_{i} \varphi_{ii}(g^{-1})\Big) = \\ &= |G|^{-1} \sum_{g\in G} \chi_\Phi(g)\,\chi_\Phi(g^{-1}) = (\chi_\Phi, \chi_\Phi)_G. \end{aligned}$$

Du fait que les caractères des représentations isomorphes coïncident, on a $(\chi_\Phi, \chi_\Psi)_G = 1$ pour $\Phi \approx \Psi$. ■

La relation (9) porte le nom de (première) *relation d'orthogonalité* pour les caractères.

COROLLAIRE. — *Soit*

$$V = V_1 \oplus \ldots \oplus V_k \tag{10}$$

une décomposition d'un G-espace complexe V en somme directe de G-sous-espaces irréductibles V_i. *Si W est un G-espace irréductible quelconque de caractère* χ_W, *le nombre de termes* V_i *dans* (10), *isomorphes à W* (*ordre de multiplicité avec lequel W figure dans le G-espace V*), *est égal à* $(\chi_V, \chi_W)_G$ *et ne dépend pas du mode de décomposition. Deux représentations* (*deux G-espaces*) *ayant même caractère sont isomorphes.*

DÉMONSTRATION. — Comme nous l'avons déjà noté (proposition (iv)), $\chi_V = \chi_{V_1} + \ldots + \chi_{V_k}$ et donc

$$(\chi_V, \chi_W)_G = (\chi_{V_1}, \chi_W)_G + \ldots + (\chi_{V_k}, \chi_W)_G.$$

En vertu du théorème 2, le second membre de cette égalité est une somme de k zéros et unités, le nombre d'unités étant égal au nombre

de G-sous-espaces V_i isomorphes à W. Or, le produit scalaire $(\chi_V, \chi_W)_G$ ne dépend aucunement de la décomposition (voir la relation de définition (8)), si bien que nous avons démontré en même temps que l'ordre de multiplicité avec lequel W figure dans le G-espace V est invariant.

Étant donné deux G-espaces V, V' ayant le même caractère $\chi = \chi_V = \chi_{V'}$, tout terme isomorphe à un G-espace irréductible donné W figure dans leur décomposition le même nombre de fois, à savoir $(\chi, \chi_W)_G$. C'est pourquoi, dans les décompositions en somme directe de termes irréductibles:

$$V = \bigoplus_{i=1}^{k} V_i, \qquad V' = \bigoplus_{j=1}^{l} V'_j$$

nous pouvons poser $l = k$, $V'_i \cong V_i$, $1 \leqslant i \leqslant k$. Par suite, les G-espaces V, V' sont eux-mêmes isomorphes. ■

Les remarques faites après la démonstration du théorème de Maschke et le corollaire du théorème 2 donnent la possibilité d'exprimer le caractère χ_Φ de toute représentation complexe linéaire (Φ, V) d'un groupe fini G sous la forme d'une combinaison linéaire à coefficients entiers:

$$\chi_\Phi = \sum_{i=1}^{s} m_i \chi_i.$$

Ici, m_i est l'ordre de multiplicité avec lequel la représentation irréductible (Φ_i, V_i) figure dans la décomposition de (Φ, V), de sorte que $\Phi_i \not\approx \Phi_j$ pour $i \neq j$. En partant de la relation d'orthogonalité (9), on peut écrire

$$(\chi_\Phi, \chi_\Phi)_G = \sum_{i=1}^{s} m_i^2. \tag{11}$$

Par suite, le *carré scalaire* $(\chi_\Phi, \chi_\Phi)_G$ *du caractère* χ_Φ *de toute représentation complexe* Φ *est toujours un nombre entier*, *égal à* 1 *si*, *et seulement si*, Φ *est une représentation irréductible*. ■

Nous avons obtenu un résultat remarquable. Les caractères ou les « traces des représentations », qui ne portent qu'une information maigre sur chaque opérateur linéaire $\Phi(g)$ pris isolément, expriment des propriétés essentielles de leur ensemble $\{\Phi(g) \mid g \in G\}$, c'est-à-dire des propriétés de la représentation Φ elle-même.

EXEMPLE 2.— Assurons-nous de l'irréductibilité sur $\mathbb{C}$ des représentations des groupes A_4, S_4 et A_5 par des rotations d'un espace à trois dimensions. A cet effet, il faut revenir au corollaire du théorème 2 du § 3 et utiliser les formules (6) et (11). La représentation Φ décrite au § 3 montre que, si σ est une permutation d'ordre q, alors $\Phi(\sigma)$ est une rotation d'angle $k\frac{2\pi}{q}$, P.G.C.D. $(k, q) = 1$, autour d'un certain axe. Aussi, les valeurs du caractère $\chi = \chi_\Phi$ se calculent-elles

directement au moyen de la formule (6):

$$\chi(\sigma) = 1 + 2\cos k\frac{2\pi}{q} = 3,\ -1,\ 0,\ 1,\ \frac{1+\sqrt{5}}{2},\ \frac{1-\sqrt{5}}{2},$$

si, respectivement, $q = 1, 2, 3, 4, 5$ ($k = \pm 1$), 5 ($k = \pm 2$). Remarquons que

$$\frac{1+\sqrt{5}}{2} = \mathrm{tr}\begin{Vmatrix} \varepsilon & 0 & 0 \\ 0 & \varepsilon^{-1} & 0 \\ 0 & 0 & 1 \end{Vmatrix} = \varepsilon + \varepsilon^{-1} + 1, \quad \frac{1-\sqrt{5}}{2} = \varepsilon^2 + \varepsilon^{-2} + 1,$$

$$\varepsilon = e^{\frac{2\pi i}{5}}.$$

Le calcul de l'ordre de la permutation σ à partir de sa décomposition en cycles indépendants a été décrit dans le corollaire 1 au théorème 4 du chapitre 4, § 2. La distribution des éléments suivant les classes de conjugaison est indiquée dans les tables (pour A_4 voir chap. 7, § 2, exercice 8, pour S_4 voir chap. 7, § 3, exercice 4, et pour A_5 voir la démonstration du théorème 5 au chap. 7, § 3). Nous reproduisons ci-dessous les mêmes tables complétées de valeurs du caractère χ:

A_4	1	3	4	4
	e	(12) (34)	(123)	(132)
χ	3	—1	0	0

S_4	1	3	6	8	6
	e	(12) (34)	(12)	(123)	(1234)
χ	3	—1	—1	0	1

A_5	1	15	20	12	12
	e	(12) (34)	(123)	(12345)	(12354)
χ	3	—1	0	$(1+\sqrt{5})/2$	$(1-\sqrt{5})/2$

Les relations

$$(\chi, \chi)_{A_4}=\frac{1}{12}\{1\cdot 3^2+3(-1)^2+4\cdot 0^2+4\cdot 0^2\}=1,$$

$$(\chi, \chi)_{S_4}=\frac{1}{24}\{1\cdot 3^2+3(-1)^2+6(-1)^2+8\cdot 0^2+6\cdot 1^2\}=1,$$

$$(\chi, \chi)_{A_5}=\frac{1}{60}\{1\cdot 3^2+15(-1)^2+20\cdot 0^2+12\left(\frac{1+\sqrt{5}}{2}\right)^2+ +12\left(\frac{1-\sqrt{5}}{2}\right)^2\right\}=1$$

montrent que la représentation Φ de caractère χ est irréductible sur $\mathbb{C}$ (voir (11)).

EXERCICES

1. Soient Φ, Ψ deux représentations complexes irréductibles d'un groupe fini G. Obtenir la généralisation du théorème 2:

$$|G|^{-1}\sum_g \chi_\Psi(hg)\overline{\chi_\Phi(g)}=\delta_{\Phi,\Psi}\frac{\chi_\Phi(h)}{\chi_\Phi(e)}.$$

Ici, h est un élément arbitraire du groupe G; $\delta_{\Phi,\Psi}=1$ ou 0 suivant que Φ et Ψ sont équivalentes ou non. (I n d i c a t i o n. Mettre les relations (4) et (5) sous la forme

$$|G|^{-1}\sum_g \psi_{jj_0}(g)\,\varphi_{i_0 i}(g^{-1})=\delta_{\Phi,\Psi}\frac{\delta_{ji}\delta_{j_0 i_0}}{\chi_\Phi(e)}.$$

Multiplier les deux membres par $\psi_{kj}(h)$ et sommer sur j, en tenant compte de l'égalité $\sum_j \psi_{kj}(h)\,\psi_{jj_0}(g)=\psi_{kj_0}(hg)$. Dans la relation ainsi obtenue

$$|G|^{-1}\sum_g \psi_{kj_0}(hg)\,\varphi_{i_0 i}(g^{-1})=\delta_{\Phi,\Psi}\frac{\psi_{ki}(h)\,\delta_{j_0 i_0}}{\chi_\Phi(e)}$$

poser $j_0=k$, $i_0=i$ et, en effectuant une sommation sur i et k, passer aux caractères.)

2. Appliquer le critère d'irréductibilité basé sur les caractères à la représentation $\Phi^{(3)}$ du groupe S_3 considéré au § 2, nº 1, exemple 1.

3. Démontrer à l'aide du lemme de Schur que toutes les représentations irréductibles sur $\mathbb{C}$ d'un groupe abélien G sont de dimension un. (I n d i c a t i o n. Soient Φ une représentation irréductible, h un élément de G. En raison de la commutativité on a $\Phi(g)\,\Phi(h)=\Phi(h)\,\Phi(g)$, $\forall g\in G$. En posant $\sigma = \Phi(h)$ dans le lemme de Schur, on obtient $\Phi(h)=\lambda_h\,\mathcal{E}$. Cela est vrai pour tout $h\in G$. Il ne reste qu'une seule possibilité pour la représentation irréductible Φ, celle d'être de dimension un.)

4. Si un groupe G possède un automorphisme τ, à toute représentation linéaire (Φ, V) de ce groupe est associée encore une représentation (Φ^τ, V) définie par $\Phi^\tau(g)=\Phi(\tau(g))$. Vérifier qu'il en est ainsi et montrer que l'irréductibilité de Φ entraîne celle de Φ^τ. En règle générale, $\Phi^\tau\approx\Phi$, mais il y a des cas où l'on obtient une nouvelle représentation. Que faut-il attendre dans le cas d'un automorphisme intérieur?

Soient $G = A_5$ et Φ la représentation considérée dans l'exemple 2. L'application $\tau : \pi \mapsto (12)\, \pi\, (12)^{-1}$ est un automorphisme (extérieur) du groupe A_5 qui permute les classes de représentants (12345) et (12354). Les ensembles des valeurs des caractères χ et χ^{τ} sont obtenus l'un à partir de l'autre par permutation de $(1 + \sqrt{5})/2$ et $(1 - \sqrt{5})/2$. Montrer que χ et χ^{τ} sont des caractères des représentations non équivalentes.

5. Soient $\Phi : G \to U(n)$, $\Psi : G \to U(n)$ deux représentations unitaires irréductibles équivalentes d'un groupe fini G. Démontrer qu'il existe une matrice unitaire U telle que $U\Phi_g U^{-1} = \Psi_g$, $\forall g \in G$. (Indication. Par hypothèse, $C\Phi_g C^{-1} = \Psi_g$ pour une certaine matrice $C = (c_{ij}) \in \mathrm{GL}(n, \mathbb{C})$. L'opération $A \mapsto A^* = {}^t\bar{A}$ appliquée à $C\Phi_g = \Psi_g C$ donne $\Phi_g^{-1} C^* = C^* \Psi_g^{-1}$, d'où $\Phi_g^{-1} C^* C = C^* C \Phi_g^{-1}$. D'après le lemme de Schur, $C^*C = \lambda E$. On a ensuite $\lambda = \sum_{k=1}^{n} | c_{ki} |^2 = \mu\bar{\mu}$, $\mu \in \mathbb{C}$, et $U = \mu^{-1} C$ est la matrice unitaire cherchée.)

§ 5. Représentations irréductibles des groupes finis

1. Nombre de représentations irréductibles.— Dans le cas des groupes finis, les considérations, qui précèdent, permettent de répondre à des questions de principe de la théorie des représentations. L'un des théorèmes fondamentaux est le suivant :

THÉORÈME 1. — *Le nombre de représentations irréductibles deux à deux non équivalentes d'un groupe fini G sur $\mathbb{C}$ est égal au nombre de ses classes des éléments conjugués.*

La démonstration de ce théorème est contenue dans les lemmes 1 et 2 si l'on remarque que le nombre r de classes de conjugaison du groupe G est interprété comme dimension de l'espace $X_{\mathbb{C}}(G)$ des fonctions centrales à valeurs complexes sur G (voir (7) du § 4). Puisque les caractères des représentations linéaires sont des fonctions centrales, ils engendrent dans $X_{\mathbb{C}}(G)$ un sous-espace linéaire d'une dimension $s \leqslant r$. D'après le théorème 2 du § 4, les caractères des représentations irréductibles forment une base orthonormée (dans la métrique $(*, *)_G$) de ce sous-espace. Par conséquent, le nombre qui nous intéresse est égal à s et n'est pas supérieur à r. Il ne reste qu'à établir l'égalité $s = r$.

LEMME 1. — *Soient Γ une fonction centrale sur un groupe fini G et (Φ, V) une représentation irréductible sur $\mathbb{C}$ de caractère χ_Φ. Alors, pour l'opérateur linéaire*

$$\Phi_\Gamma = \sum_{h \in G} \bar{\Gamma}(h)\, \Phi(h) : V \to V$$

on a $\Phi_\Gamma = \lambda \mathcal{E}$, où

$$\lambda = \frac{|G|}{\chi_\Phi(e)} (\chi_\Phi, \Gamma)_G$$

($\bar{\Gamma}$ est une fonction centrale définie par l'égalité $\bar{\Gamma}(g) = \overline{\Gamma(g)}$).

DÉMONSTRATION. — Puisque Γ est une fonction centrale, on a

$$\Phi(g)\,\Phi_\Gamma\Phi(g)^{-1} = \sum_{h\in G} \overline{\Gamma}(h)\,\Phi(g)\,\Phi(h)\,\Phi(g^{-1}) =$$

$$= \sum_{h\in G} \overline{\Gamma}(ghg^{-1})\,\Phi(ghg^{-1}) = \sum_{t\in G} \overline{\Gamma}(t)\,\Phi(t) = \Phi_\Gamma.$$

Ainsi, $\Phi_\Gamma\Phi(g) = \Phi(g)\,\Phi_\Gamma$, $\forall g \in G$. Le lemme de Schur (théorème 1 du § 4) appliqué au cas $\sigma = \Phi_\Gamma$ montre que $\Phi_\Gamma = \lambda\mathscr{E}$. En calculant la trace des opérateurs intervenant aux deux membres de cette égalité, on trouve

$$\lambda\chi_\Phi(e) = \lambda \dim V = \operatorname{tr}\lambda\mathscr{E} = \operatorname{tr}\Phi_\Gamma = \sum_{h\in G} \overline{\Gamma}(h)\operatorname{tr}\Phi(h) =$$

$$= |G|\left\{|G|^{-1}\sum_{h\in G} \chi_\Phi(h)\,\overline{\Gamma(h)}\right\} = |G|\,(\chi_\Phi, \Gamma)_G. \quad \blacksquare$$

LEMME 2. — *Les caractères $\chi_1, \ldots, \chi_s$ de toutes les représentations irréductibles, deux à deux non équivalentes d'un groupe G sur $\mathbb{C}$ forment une base orthonormée de l'espace $X_\mathbb{C}(G)$.*

DÉMONSTRATION. — D'après le théorème 2 du § 4, le système $\chi_1, \ldots, \chi_s$ est orthonormé et peut être inclus dans la base orthonormée de l'espace $X_\mathbb{C}(G)$. Soit Γ une fonction centrale arbitraire orthogonale à tous les χ_i: $(\chi_i, \Gamma)_G = 0$. Alors, en vertu du lemme 1, l'opérateur linéaire $\Phi_\Gamma^{(i)}$ correspondant à la représentation $\Phi^{(i)}$ de caractère χ_i est nul.

D'après le théorème de Maschke, toute représentation complexe Φ peut être décomposée en somme directe

$$\Phi = m_1\Phi^{(1)} \dotplus \ldots \dotplus m_s\Phi^{(s)}$$

avec certains ordres de multiplicité $m_1, \ldots, m_s$. Conformément à cette décomposition, on a

$$\Phi_\Gamma = m_1\Phi_\Gamma^{(1)} + \ldots + m_s\Phi_\Gamma^{(s)} = 0$$

pour l'opérateur Φ_Γ défini par la relation

$$\Phi_\Gamma = \sum_{h\in G} \overline{\Gamma}(h)\,\Phi(h).$$

En particulier, cela est valable pour l'opérateur linéaire ρ_Γ, où ρ est la représentation régulière (voir § 1, exemple 5). Or, dans un tel cas on aura (en désignant pour un instant l'élément unité du groupe G par le symbole 1, pour éviter d'employer la notation e_e)

$$0 = \rho_\Gamma(e_1) = \sum_{h\in G} \overline{\Gamma}(h)\,\rho(h)\,(e_1) = \sum_{h\in G} \overline{\Gamma}(h)\,e_h \Rightarrow \overline{\Gamma}(h) = 0, \quad \forall h \in G,$$

d'où $\overline{\Gamma} = 0$ et par conséquent $\Gamma = 0$. $\blacksquare$

EXEMPLE.— Le théorème 1 appliqué au groupe symétrique S_3 affirme que ce groupe possède exactement trois représentations complexes irréductibles. Il n'est pas besoin de les chercher: la table donnée à la fin du nº 1 du § 2 fournit toute l'information nécessaire. Remarquons, entre autres, que les carrés des degrés des représentations $\Phi^{(1)}$, $\Phi^{(2)}$, $\Phi^{(3)}$ vérifient la relation $1^2 + 1^2 + 2^2 = 6 = |S_3|$. Nous allons voir maintenant qu'une relation analogue est aussi vérifiée dans le cas général.

2. Degrés des représentations irréductibles.— Considérons de plus près la représentation régulière $(\rho, \langle e_g \mid g \in G\rangle_{\mathbb{C}})$. (Désignons par R_h la matrice de l'opérateur linéaire $\rho(h)$ dans la base donnée $\{e_g \mid g \in G\}$. Puisque $\rho(h)e_g = e_{hg}$, tous les éléments diagonaux de la matrice R_h sont nuls pour $h \neq e$, et $\operatorname{tr} R_h = 0$. Par suite

$$\chi_\rho(e) = |G|, \quad \chi_\rho(h) = 0, \quad \forall h \neq e. \tag{1}$$

Maintenant soit (Φ, V) une représentation irréductible arbitraire du groupe G sur $\mathbb{C}$. Comme le montre le corollaire au théorème 2 du § 4, l'ordre de multiplicité, avec lequel Φ figure dans ρ, est égal au produit scalaire $(\chi_\rho, \chi_\Phi)_G$. D'après (1) on a

$$(\chi_\rho, \chi_\Phi)_G = |G|^{-1}\sum_{h\in G} \chi_\rho(h)\overline{\chi_\Phi(h)} = |G|^{-1}\chi_\rho(e)\overline{\chi_\Phi(e)} = $$
$$= |G|^{-1}|G|\chi_\Phi(e) = \dim V. \tag{2}$$

On voit que toute représentation irréductible (considérée à une équivalence près) apparaît dans la représentation régulière avec un ordre de multiplicité égal à son degré. Suivant le théorème 1, il existe r représentations irréductibles deux à deux non équivalentes

$$\Phi^{(1)}, \Phi^{(2)}, \ldots, \Phi^{(r)}$$

(r est le nombre de classes de conjugaison du groupe G) auxquelles correspondent les caractères

$$\chi_1, \chi_2, \ldots, \chi_r; \quad \chi_i = \chi_{\Phi^{(i)}},$$

de degrés

$$n_1, n_2, \ldots, n_r; \; n_i = \chi_i(e).$$

Pour $\Phi^{(1)}$ on prend ordinairement la représentation unité, si bien que $\chi_1(g) = 1$, $\forall g \in G$. La relation (2) montre que

$$\rho = n_1\Phi^{(1)} + \ldots + n_r\Phi^{(r)},$$

d'où

$$\chi_\rho = n_1\chi_1 + \ldots + n_r\chi_r.$$

En particulier,

$$|G| = \chi_\rho(e) = n_1\chi_1(e) + \ldots + n_r\chi_r(e) = n_1^2 + \ldots + n_r^2.$$

Nous sommes arrivés au théorème suivant:

THÉORÈME 2. — *Toute représentation irréductible $\Phi^{(i)}$ apparaît dans la décomposition de la représentation régulière ρ avec un ordre*

de multiplicité égal à son degré n_i. L'ordre $|G|$ d'un groupe fini G et les degrés $n_1, \ldots, n_r$ de toutes ses représentations irréductibles non équivalentes sont liés par la relation

$$\sum_{i=1}^{r} n_i^2 = |G|. \quad \blacksquare \tag{3}$$

Dans le cas des groupes de petit ordre, la relation élégante (3) est suffisante pour trouver tous les degrés $n_1, \ldots, n_r$, alors que dans le cas général, des considérations supplémentaires sont évidemment nécessaires.

Les renseignements sur les caractères des représentations irréductibles (ou bref, sur les caractères irréductibles) peuvent être commodément écrits sous la forme d'une table

	e	g_2	g_3	. . . .	g_r
χ_1	n_1	$\chi_1(g_2)$	$\chi_1(g_3)$	. . . . . .	$\chi_1(g_r)$
χ_2	n_2	$\chi_2(g_2)$	$\chi_2(g_3)$	. . . . . .	$\chi_2(g_r)$
. . .	. .	. . . .	. . . .	. . . . . .	. . . . . .
χ_r	n_r	$\chi_r(g_2)$	$\chi_r(g_3)$	. . . . . .	$\chi_r(g_r)$

que l'on appelle *table de caractères*. Sa ligne supérieure indique les représentants de toutes les r classes de conjugaison g_i^G du groupe G. Indiquons à titre d'exemple la table de caractères pour le groupe S_3:

	e	(12)	(123)
χ_1	1	1	1
χ_2	1	—1	1
χ_3	2	0	—1

(comparer avec la table donnée à la fin du n° 1 du § 2). Désignons comme toujours par le symbole $C(g) = C_G(g)$ le centralisateur dans le groupe G de l'élément $g \in G$. Nous savons que $|C(g)|\,|g^G| = |G|$ (voir chap. 7, § 2, n° 2). Aussi, la relation (9) du § 4 (première relation d'orthogonalité) mise sous la forme

$$\sum_{j=1}^{r} \frac{\chi_i(g_j)}{\sqrt{|C(g_j)|}} \, \frac{\overline{\chi_k(g_j)}}{\sqrt{|C(g_j)|}} = \frac{1}{|G|} \sum_{j=1}^{r} \frac{|G|}{|C(g_j)|} \chi_i(g_j)\overline{\chi_k(g_j)} =$$

$$= \frac{1}{|G|} \sum_{j=1}^{r} |g_j^G|\, \chi_i(g_j)\overline{\chi_k(g_j)} = \frac{1}{|G|} \sum_{g \in G} \chi_i(g)\overline{\chi_k(g)} = (\chi_i, \chi_k)_G = \delta_{ik}$$

signifie-t-elle que la matrice $r \times r$

$$M = \left(\frac{\chi_i(g_j)}{\sqrt{|C(g_j)|}}\right)$$

est unitaire suivant les lignes. Or, si une matrice est unitaire suivant les lignes, elle l'est suivant les colonnes $(M \cdot {}^t\overline{M} = E = {}^t\overline{M} \cdot M)$, si bien que

$$\sum_i \frac{\chi_i(g_j)}{\sqrt{|C(g_j)|}} \frac{\overline{\chi_i(g_k)}}{\sqrt{|C(g_k)|}} = \delta_{jk},$$

ou encore, sous une forme plus détaillée:

$$\sum_{i=1}^{r} \chi_i(g)\overline{\chi_i(h)} = \begin{cases} 0 & \text{si } g \text{ et } h \text{ ne sont pas conjugués,} \\ |C_G(g)| & \text{dans le cas contraire.} \end{cases} \quad (4)$$ ■

La relation (4) porte le nom de *deuxième relation d'orthogonalité* pour les caractères.

3. Représentations des groupes abéliens.— La description des représentations irréductibles des groupes cycliques, donnée dans l'exemple 6 du § 1, admet une généralisation naturelle suivante:

THÉORÈME 3. — *Toute représentation irréductible d'un groupe abélien fini A sur $\mathbb{C}$ est de degré 1. Le nombre de telles représentations deux à deux non équivalentes est égal à l'ordre $|A|$. Réciproquement, si chaque représentation irréductible d'un groupe fini A est de degré 1, A est un groupe abélien.*

DÉMONSTRATION. — Comme le nombre r de classes des éléments conjugués du groupe abélien A coïncide avec son ordre, les deux premières assertions découlent du théorème 2 (voir aussi § 4, exercice 3). En posant maintenant dans la relation (3) tous les n_i égaux à 1, on obtient $r = |A|$, ce qui est équivalent à la commutativité du groupe. ■

DÉFINITION. — *Soit A un groupe abélien. L'ensemble*

$$\hat{A} = \mathrm{Hom}(A, \mathbb{C}^*)$$

des homomorphismes du groupe A dans le groupe multiplicatif $\mathbb{C}^*$ du corps des nombres complexes muni de l'opération de multiplication définie par

$$(\chi_1\chi_2)(a) = \chi_1(a)\chi_2(a)$$

$(\chi_i \in \hat{A},\ a \in A)$ *s'appelle groupe des caractères du groupe A sur* $\mathbb{C}(\chi^{-1} = \overline{\chi})$.

THÉORÈME 4. — *Les groupes A et $\hat{A}$ sont isomorphes.*

DÉMONSTRATION. — Du théorème 3 nous déduisons qu'en tout cas $|A| = |\hat{A}|$. D'après les résultats du chapitre 7, § 5, le groupe A admet une décomposition

$$A = A_1 \times A_2 \times \ldots \times A_k$$

en produit direct des groupes cycliques $A_i = \langle a_i \rangle$ (peu importe qu'ils soient primaires ou non; nous choisissons la notation multiplicative de la loi de multiplication dans A). Si $|A_i| = s_i$ et ε_i est une racine primitive s_i-ième de l'unité, à tout élément $a = a_1^{t_1} a_2^{t_2} \ldots a_k^{t_k}$ de A correspond un caractère $\chi_a \in \hat{A}$ défini par la relation

$$\chi_a (a_1^{r_1} a_2^{r_2} \ldots a_k^{r_k}) = \varepsilon_1^{r_1 t_1} \varepsilon_2^{r_2 t_2} \ldots \varepsilon_k^{r_k t_k}.$$

Il est évident que $\chi_a \chi_{a'} = \chi_{aa'}$ (voir définition). Si

$$a = a_1^{t_1} a_2^{t_2} \ldots a_k^{t_k} \neq a_1^{t_1'} a_2^{t_2'} \ldots a_k^{t_k'} = a',$$

il existe un indice i, tel que $t_i \neq t_i'$. Alors

$$\chi_a (a_i) = \varepsilon_i^{t_i} \neq \varepsilon_i^{t_i'} = \chi_{a'} (a_i).$$

Par conséquent, tous les caractères χ_a sont deux à deux distincts et l'application $a \mapsto \chi_a$ établit l'isomorphisme cherché entre A et $\hat{A}$. ∎

La méthode, utilisée pour démontrer le théorème 4, fournit manifestement une construction explicite de toutes les représentations irréductibles d'un groupe abélien.

EXEMPLE. — Soient V_{2^n} un groupe abélien élémentaire d'ordre 2^n et χ son caractère complexe irréductible différent du caractère unité, c'est-à-dire $\chi(a) \neq 1$ pour un certain $a \in V_{2^n}$. Alors, Ker $\chi = B \cong V_{2^{n-1}}$ et il existe une décomposition $V_{2^n} = B \cup aB$ en classes suivant B, de sorte que

$$\chi(a^i b) = (-1)^i, \qquad i = 0, 1.$$

En particulier, le groupe à quatre éléments (groupe de Klein) V_4, dont les représentations ont été mentionnées au chapitre 1, § 2, problème 2, possède une table de caractères suivante

	e	a	b	ab
χ_1	1	1	1	1
χ_2	1	-1	1	-1
χ_3	1	1	-1	-1
χ_4	1	-1	-1	1

Les résultats relatifs aux représentations des groupes abéliens permettent d'obtenir aussi une certaine information sur les représentations des groupes finis arbitraires.

THÉORÈME 5. — *Les représentations de degré* 1 *d'un groupe fini* G *sur* $\mathbb{C}$ *sont en bijection avec les représentations irréductibles du groupe quotient* G/G' (G' *est le sous-groupe dérivé du groupe* G). *Leur nombre est égal à l'indice* $(G : G')$.

DÉMONSTRATION. — Faisons d'abord une remarque générale. Soient G un groupe quelconque et K son sous-groupe distingué. Si Φ est une représentation du groupe G avec le noyau $\operatorname{Ker}\Phi \supset K$, on peut définir une représentation du groupe quotient G/K en posant

$$\overline{\Phi}(gK) = \Phi(g), \qquad g \in G.$$

Il est évident que cette définition est correcte (voir chap. 7, § 3, démonstration du théorème 1). On a ensuite $\operatorname{Ker}\overline{\Phi} = \operatorname{Ker}\Phi/K$. En particulier, pour $K = \operatorname{Ker}\Phi$, on obtient une représentation exacte $\overline{\Phi}$.

Réciproquement, toute représentation linéaire Ψ d'un groupe H induit une représentation Φ du groupe G, admettant un épimorphisme $\pi : G \rightarrow H$. Il suffit de poser

$$\Phi(g) = \Psi(\pi(g)).$$

Puisque π est un épimorphisme, $\Phi(G) = \Psi(H)$ et Φ, Ψ sont simultanément réductibles ou irréductibles. D'après le théorème de correspondance (chap. 7, § 3, théorème 3), $\operatorname{Ker}\Phi = \pi^{-1}(\operatorname{Ker}\Psi)$. A toute représentation Φ de dimension un d'un groupe G est associé un groupe abélien (plus exactement, cyclique) $\operatorname{Im}\Phi$, si bien que $\operatorname{Ker}\Phi \supset G'$. La démonstration du théorème est maintenant obtenue par une simple réunion du théorème 3, de la remarque faite plus haut et du théorème 4 du chapitre 7, § 3. ■

4. Représentations de certains groupes spéciaux.— Bien qu'il suffise en principe de décomposer la représentation régulière d'un groupe fini G pour obtenir toutes ses représentations irréductibles (théorème 2), en pratique, on aura de grandes difficultés à utiliser ce procédé et on est donc conduit à chercher des voies de détour. Généralement, il s'avère plus facile de composer d'abord la table de caractères et de construire ensuite les représentations elles-mêmes (voir à ce propos chap. 9, § 1). D'ailleurs dans les exemples relativement simples que nous donnons ci-dessous, il n'est pas besoin d'avoir recours à un artifice quelconque.

EXEMPLE 1. — Soit G un groupe 2-transitif de permutations, opérant sur l'ensemble $\Omega = \{1, 2, \ldots, n\}$, $n > 2$ (voir chap. 7, § 2, exemple 3). Soit ensuite Φ la représentation naturelle du groupe G sur l'espace $V = \langle e_1, e_2, \ldots, e_n\rangle$ avec l'opération $\Phi(g)\, e_i = e_{g(i)}$ (voir § 1, exemple 5). On comprend aisément que la valeur de $\chi_\Phi(g)$ coïncide avec le nombre $N(g)$ de points $i \in \Omega$ (= de vecteurs de base e_i) restant fixes lors de l'opération de g. D'après le

théorème 3 du chapitre 7, § 2, on a

$$\sum_{g\in G} \chi_\Phi(g)\overline{\chi_\Phi(g)} = \sum_{g\in G} \chi_\Phi(g)^2 = \sum_{g\in G} N(g)^2 = 2\,|G|,$$

ce qui peut évidemment se mettre sous la forme

$$(\chi_\Phi, \chi_\Phi)_G = 2. \tag{5}$$

En rapprochant la relation (5) de la relation (11) du § 4, nous arrivons à la conclusion que Φ est somme directe de deux représentations irréductibles ($2 = 1 + 1$ est l'unique expression du nombre 2 sous forme de la somme des carrés des entiers naturels). Or, nous savons aussi que $\Phi = \Phi^{(1)} \dot{+} \Psi$, où $(\Phi^{(1)}, U)$ est la représentation unité et Ψ est une représentation de dimension $(n - 1)$ opérant sur l'espace $W = \langle e_1 - e_n, e_2 - e_n, \ldots, e_{n-1} - e_n \rangle$. Si la décomposition $V = U \oplus W$ pouvait être prolongée grâce à une décomposition de W, le nombre de termes irréductibles serait supérieur à deux. Ainsi, on a l'assertion non triviale suivante:

La représentation linéaire naturelle (Φ, V) *d'un groupe 2-transitif de permutations* G *sur le corps* $\mathbb{C}$ *est la somme de la représentation unité et d'une représentation irréductible.* ■

En particulier, chacun des groupes S_n, $n > 2$; A_n, $n > 3$, *possède une représentation irréductible* Ψ *sur* $\mathbb{C}$ *de degré* $n - 1$ *et de caractère* χ_Ψ *calculé par la formule*

$$\chi_\Psi(g) = N(g) - 1. \;■ \tag{6}$$

Comme il a été montré sur l'exemple du groupe S_3 (§ 2, n° 1, exemple 1), les matrices Ψ_g se déterminent assez facilement. Pour pouvoir calculer les valeurs de $\chi_\Psi(g)$ à l'aide de la formule (6), il suffit de connaître la structure des cycles de la permutation g. Nous en donnons ci-dessous une petite illustration:

A_4	e	(12) (34)	(123)	(132)
χ_Ψ	3	—1	0	0

S_4	e	(12) (34)	(12)	(123)	(1234)
χ_Ψ	3	—1	1	0	—1

A_5	e	(12) (34)	(123)	(12345)	(12354)
χ_Ψ	4	0	1	—1	—1

EXEMPLE 2. *Représentations irréductibles du groupe alterné A_4.* — Révisons des faits déjà connus. Le groupe A_4 possède quatre classes d'éléments conjugués. Les représentants des classes et leurs puissances sont indiqués dans les deux lignes supérieures de la table ci-dessous :

	1	3	4	4
	e	(12) (34)	(123)	(132)
χ_1	1	1	1	1
χ_2	1	1	ε	ε^{-1}
χ_3	1	1	ε^{-1}	ε
χ_4	3	—1	0	0

Le sous-groupe dérivé $A_4' = \{e, (12)(34), (13)(24), (14)(23)\} = V_4$ a dans A_4 l'indice 3, et le groupe A_4 possède donc trois représentations de dimension un $\Phi^{(1)} = \chi_1$, $\Phi^{(2)} = \chi_2$, $\Phi^{(3)} = \chi_3$ (avec le noyau A_4' et $\varepsilon^3 = 1$, $\varepsilon \neq 1$) et une représentation de dimension trois $\Phi^{(4)}$ $(12 = 1^2 + 1^2 + 1^2 + 3^2)$. En comparant les tables pour A_4 données dans l'exemple 1 et dans l'exemple 2 du § 4, on s'assure que la représentation $\Phi^{(4)}$ de caractère χ_4 est équivalente à la représentation Φ du groupe A_4 par les rotations (groupe du tétraèdre) et la représentation Ψ liée à la 2-transitivité du groupe A_4.

EXEMPLE 3. *Représentations irréductibles du groupe symétrique S_4.* — Deux lignes supérieures de la table sont prises au chapitre 7,

	1	3	6	8	6
	e	(12) (34)	(12)	(123)	(1234)
χ_1	1	1	1	1	1
χ_2	1	1	—1	1	—1
χ_3	2	2	0	—1	0
χ_4	3	—1	—1	0	1
χ_5	3	—1	1	0	—1

§ 3, exercice 4. $\Phi^{(1)} = \chi_1$ est la représentation unité. La représentation $\Phi^{(2)} = \chi_2$ est définie par la signature des permutations de S_4. Puisque $(S_4 : S_4') = 2$ (exemple du chap. 7, § 3, n° 2), d'autres

représentations de dimension un n'existent pas. La représentation de dimension deux $\Phi^{(3)}$ de caractère χ_3 et de noyau $V_4 \lhd S_4$ est obtenue en partant des considérations exposées au cours de la démonstration du théorème 5 et au chapitre 7, § 3, n° 1, exemple 2. La représentation $\Phi^{(4)}$ de caractère χ_4 correspond aux rotations d'un cube (voir table pour S_4 au § 4, exemple 2). La représentation $\Phi^{(5)} = \Psi$ de caractère χ_5 (voir table dans l'exemple 1) est liée à la 2-transitivité du groupe S_4. Elle est aussi équivalente à la représentation associée à toutes les transformations de symétrie du tétraèdre $\triangle_4$ (rotations + réflexions; ce sont justement ces transformations qui ont de l'importance pour la description des oscillations d'une molécule de phosphore (chap. 1, § 2, problème 2)).

EXEMPLE 4. *Représentations irréductibles du groupe quaternionien* Q_8. — Tout ce qui concerne le groupe Q_8 a été dit au chapitre 7, § 3, n° 5, exemple 2. On y a également donné (sans l'appeler de son nom) la représentation irréductible $\Phi^{(5)}$ de dimension deux et de caractère χ_5.

	1	1	2	2	2
	e	a^2	a	b	ab
χ_1	1	1	1	1	1
χ_2	1	1	—1	—1	1
χ_3	1	1	—1	1	—1
χ_4	1	1	1	—1	—1
χ_5	2	—2	0	0	0

Les quatre représentations de dimension un ont pour noyau le sous-groupe dérivé $\langle a^2 \rangle$ et sont déterminées à l'aide de la table donnée dans l'exemple du n° 3.

EXERCICES

1. Etant donné la décomposition $\Gamma_i = \sum_j t_{ij}\chi_j$ de la fonction centrale de base Γ_i (voir § 4, relation (7)), obtenir la relation (4) en exprimant les coefficients $t_{ij} = (\Gamma_i, \chi_j)_G$ de cette décomposition en fonction des caractères irréductibles.

2. Vérifier (et se rappeler l'isomorphisme entre l'espace vectoriel V et son dual V^* des fonctions linéaires) que l'application $\tau : A \to \hat{\hat{A}}$ définie par la condition

$$a^\tau(\chi) = \chi(a)$$

détermine l'isomorphisme du groupe abélien A sur $\hat{\hat{A}}$. (Indication. De $a^\tau(\chi_1\chi_2) = a^\tau(\chi_1)a^\tau(\chi_2)$, il s'ensuit que a^τ est le caractère du groupe $\hat{A}$. Puisque $(aa')^\tau = a^\tau(a')^\tau$, τ est un homomorphisme de A dans $\hat{\hat{A}}$. On a ensuite

$$\operatorname{Ker} \tau = \{a \in A \mid a^\tau(\chi) = \chi(a) = 1, \quad \forall \chi \in \hat{A}\} \Rightarrow \operatorname{Ker} \tau = e,$$

et $|\hat{\hat{A}}| = |\hat{A}| = |A|$. Par conséquent τ est un isomorphisme.)

Cet exercice joint au théorème 4 établit une partie de la loi dite *loi de dualité pour les groupes abéliens finis*. Une loi de dualité analogue mais beaucoup plus profonde, relative aux groupes abéliens topologiques, qui conduit à des conséquences bien importantes, a été établie aux années trente par L. S. Pontriaguine.

3. Démontrer que, si un groupe abélien fini A admet une représentation complexe irréductible exacte, alors A est un groupe cyclique.

4. Soient A un groupe abélien fini et B son sous-groupe. Démontrer que tout caractère du groupe B est prolongeable en un caractère du groupe A et que le nombre de tels prolongements est égal à l'indice $(A:B)$.

5. Justifier la proposition qui précède les parenthèses finales dans l'exemple 3 du nº 4.

6. Quelle est la moyenne $\frac{1}{|G|}\sum_g \chi(g)$ des valeurs du caractère complexe χ sur les éléments d'un groupe fini G?

7. Réunir les tables relatives au groupe A_5 données dans de différents endroits (voir § 4, nº 2, exemple 2; § 4, exercice 4; § 5, exemple 1) en une table récapitulative de caractères

	1	15	20	12	12
	e	(12) (34)	(123)	(12345)	(12354)
χ_1	1	1	1	1	1
χ_2	3	-1	0	$\frac{1}{2}(1+\sqrt{5})$	$\frac{1}{2}(1-\sqrt{5})$
χ_3	3	-1	0	$\frac{1}{2}(1-\sqrt{5})$	$\frac{1}{2}(1+\sqrt{5})$
χ_4	4	0	1	-1	-1
χ_5	*	*	*	*	*

Donner la description des représentations irréductibles de caractères χ_1, χ_2, χ_3, χ_4. Remplir la dernière ligne de la table en utilisant la deuxième relation d'orthogonalité (4) pour les caractères. (Réponse: 5 1, -1, 0, 0.)

8. Soient $P = \{A^iB^jC^k;\ 0 \leqslant i, j, k \leqslant p-1\}$ le groupe d'ordre p^3 que nous avons considéré au chapitre 7, § 2, exercice 3; $V = \langle e_0, e_1, \ldots, e_{p-1}\rangle_{\mathbb{C}}$ un espace vectoriel complexe de dimension p; ε une racine primitive p-ième de l'unité; $\mathcal{A}$, $\mathcal{B}_k$, $\mathcal{C}_k$ des opérateurs linéaires sur V définis par les relations

$$\mathcal{A}e_i = e_{i+1},\ \mathcal{B}_k e_i = \varepsilon^{-ki}e_i,\ \mathcal{C}_k e_i = \varepsilon^k e_i,\ 0 \leqslant i \leqslant p-1$$

(les indices inférieurs des éléments de base sont pris modulo p).

Montrer que l'application

$$\Phi^{(k)} : A \mapsto \mathcal{A},\ B \mapsto \mathcal{B}_k,\ C \mapsto \mathcal{C}_k$$

définit une représentation linéaire irréductible du groupe P. Les représentations $\Phi^{(1)}, \ldots, \Phi^{(p-1)}$ sont deux à deux non équivalentes. Jointes à p^2 représentations de dimension un (p^2 est l'indice du sous-groupe dérivé $P' = \langle C\rangle$ de P), elles sont les seules représentations complexes irréductibles du groupe P.

9. Compléter de calculs le raisonnement suivant. Soit $D_n = \langle a, b \mid a^n = e, b^2 = e, bab^{-1} = a^{-1}\rangle$ le groupe diédral d'ordre $2n$ dont les propriétés (y compris la description des classes d'éléments conjugués) sont indiquées au chapitre 7, § 3, nº 5, exemple 1. Puisque $\langle a\rangle \lhd D_n$, les applications $a \mapsto 1$, $b \mapsto 1$ et $a \mapsto 1$, $b \mapsto -1$ définissent deux représentations de dimension un. Soit $\varepsilon = e^{\frac{2\pi i}{n}}$ une racine primitive n-ième de l'unité. Alors, l'application

$$\Phi^{(j)} : a \mapsto \begin{Vmatrix} \varepsilon^j & 0 \\ 0 & \varepsilon^{-j} \end{Vmatrix}, \quad b \mapsto \begin{Vmatrix} 0 & 1 \\ 1 & 0 \end{Vmatrix}$$

définit une représentation de degré 2. La représentation $\Phi^{(j)}$ est irréductible pour $j = 1, 2, \ldots, \left[\frac{n-1}{2}\right]$ ($[\alpha]$ est la partie entière du nombre réel α). Pour $n = 2m$, la représentation $\Phi^{(m)}$ se décompose en somme directe de deux représentations de dimension un: $a \mapsto -1$, $b \mapsto 1$ et $a \mapsto -1$, $b \mapsto -1$. Cela s'accorde avec le fait que le sous-groupe dérivé D'_{2m} a l'indice 4 dans D_{2m}, et $D_{2m}/D'_{2m} \cong \cong Z_2 \times Z_2$. Toutes les représentations indiquées sont irréductibles; elles forment un ensemble complet des représentations complexes irréductibles du groupe diédral. Trouver la réalisation réelle des représentations $\Phi^{(j)}$. Indiquer sous forme explicite l'isomorphisme (l'équivalence) $\Phi^{(j)} \approx \Phi^{(k)}$, $k > m$, $j \leqslant m$.

10. *Groupes cristallographiques* (pour le problème 2 du chap. 1, § 2). Soient E un espace euclidien de dimension n et V un espace vectoriel muni d'un produit scalaire euclidien et associé à E. A tout déplacement d de l'espace E correspond une transformation linéaire orthogonale $\bar{d} \subset O(n)$, si bien que l'on a $\overline{d_1 d_2} = \bar{d}_1 \bar{d}_2$. On dit que le groupe D de déplacements de l'espace est un groupe *cristallographique* si la D-orbite d'un point arbitraire est discrète (n'a pas de points limites) et s'il existe un ensemble compact $M \subset E$ tel que $D(M) = \bigcup_{d \in D} d(M) = E$. Suivant le théorème de Schönflies-Bieberbache, le groupe cristallographique D contient n translations affines indépendantes qui engendrent dans D un sous-groupe distingué L, et $\bar{D} \cong D/L$ est un groupe fini (un groupe cristallographique ponctuel). Pour $n = 3$, le nombre total de groupes cristallographiques ponctuels géométriquement différents est égal à 32. Parmi ces groupes doivent, évidemment, se trouver des groupes contenant des réflexions (des déplacements non propres). Les propriétés qui définissent un groupe cristallographique entraînent que toute rotation propre de $\bar{D}$ se représente par une matrice semblable à

$$A = \begin{Vmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{Vmatrix},$$

avec $\operatorname{tr} A = 1 + 2\cos\theta \in \mathbb{Z}$. En s'appuyant sur le théorème 2 du § 3 et sur la considération qui vient d'être développée, montrer que pour $n = 3$ les seuls groupes cristallographiques ponctuels ne contenant pas de réflexions seront les groupes cycliques C_1, C_2, C_3, C_4, C_6, les groupes diédraux D_2, D_3, D_4, D_6, le groupe du tétraèdre $\mathbf{T}$ et le groupe du cube (de l'octaèdre) $\mathbf{O}$.

§ 6. Représentations des groupes SU (2) et SO (3)

Une partie de la pensée « physique » est constituée par des images concrètes liées aux représentations du groupe SO (3). Du point de vue mathématique, l'opération du groupe SO (3), qui reflète la symétrie de nombreux problèmes de physique, est en particulier intéressante par le fait qu'elle induit une opération sur l'espace des solutions de l'équation $\Delta f = 0$, où $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ est l'opérateur de dérivation de Laplace. Un analogue de dimension deux de ce problème a été considéré au début du chapitre (problème 1).

Tout élément du groupe SO (3) est produit de plusieurs opérateurs B_φ, C_θ de la forme (1) du chapitre 7, § 1. Or, B_φ n'opère pas sur z, et C_θ sur x. Aussi, l'invariance de l'équation $\Delta f = 0$ par rapport à B_φ et C_θ découle-t-elle des calculs qui ont été effectués dans le cas de dimension deux. Nous arrivons à la conclusion que l'équation $\Delta f = 0$ est invariante par tout le groupe SO (3) ou ce qui revient au même

$$\Delta f = 0 \Rightarrow \Delta (\Phi_g f) = 0, \quad \forall g \in \mathrm{SO}\ (3),$$

où $\Phi_g f$ est une fonction définie par la relation

$$(\Phi_g f)(x, y, z) = f(g^{-1}(x), g^{-1}(y), g^{-1}(z)). \tag{1}$$

D'après la condition par la transformation orthogonale g^{-1} de matrice $(a_{ij})_1^3$, la colonne de nouvelles variables est de la forme

$$\left\| \begin{matrix} g^{-1}(x) \\ g^{-1}(y) \\ g^{-1}(z) \end{matrix} \right\| = \left\| \begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{matrix} \right\| \left\| \begin{matrix} x \\ y \\ z \end{matrix} \right\|.$$

Suivant (1) on a

$$\begin{aligned}(\Phi_g (\Phi_h f))(x, y, z) &= (\Phi_h f)(g^{-1}(x), g^{-1}(y), g^{-1}(z)) = \\ &= f(h^{-1}(g^{-1}(x)), h^{-1}(g^{-1}(y)), h^{-1}(g^{-1}(z))) = \\ &= f((gh)^{-1}(x), (gh)^{-1}(y), (gh)^{-1}(z)) = (\Phi_{gh} f)(x, y, z).\end{aligned}$$

Donc,

$$\Phi_g \Phi_h = \Phi_{gh},$$

c'est-à-dire les opérateurs linéaires Φ_g, $g \in$ SO (3), agissent sur les fonctions de manière que l'application $\Phi: g \mapsto \Phi_g$ est une représentation du groupe SO (3). Ce procédé, bien naturel, de construction des représentations (que nous avons au fait déjà utilisé lors de l'examen des fonctions symétriques sur lesquelles opère le groupe S_n) est en principe applicable à une large classe de groupes et se range parmi les méthodes typiques de l'Analyse fonctionnelle. Il s'agit seulement de choisir, en partant des conditions concrètes, l'espace convenable de fonctions et de le décomposer ensuite en

sous-espaces invariants irréductibles (ce problème relève de l'Analyse harmonique).

Dans le cas du groupe SO (3), lorsque toutes les représentations irréductibles sont de dimension finie (ce fait est commun aux groupes compacts que nous ne considérons pas dans notre exposé), on prend pour fonctions les polynômes homogènes

$$f(x, y, z) = \sum_{s,t} a_{s,t} x^s y^t z^{m-s-t}$$

de degré fixe m ($m = 1, 2, 3, \ldots$). Ces derniers forment un espace P_m de dimension $\binom{m+2}{2}$ (voir chap. 5, § 2, exercice 4). Puisque $\Delta f \in P_{m-2}$, la condition $\Delta f = 0$ est équivalente à $\binom{m}{2}$ conditions linéaires pour les coefficients $a_{s,t}$. Les solutions $f \in P_m$ de l'équation $\Delta f = 0$ sont appelées *polynômes harmoniques* homogènes de degré m. L'opérateur Δ étant linéaire, ils forment un sous-espace H_m de dimension $\binom{m+2}{2} - \binom{m}{2} = 2m + 1$ (dans notre cas, de dimension $\geqslant 2m + 1$, mais en réalité il y a une égalité). Du fait de ce qui précède, H_m est invariant par l'opération $\Phi = \Phi^{(m)}$ du groupe SO (3). Il s'avère qu'a lieu le théorème suivant: *l'espace H_m de la représentation $\Phi^{(m)}$ est irréductible sur $\mathbb{C}$; toute représentation irréductible sur $\mathbb{C}$ du groupe* SO (3) *est équivalente à l'une des représentations* ($\Phi^{(m)}$, H_m) *de dimension impaire* $2m + 1$. Au lieu de démontrer ce théorème, nous nous contenterons de ce qui vient d'être dit et passerons au groupe SU (2) pour lequel il est un peu plus facile d'obtenir la famille de représentations irréductibles. Etant donné qu'il existe l'épimorphisme naturel SU (2) $\to$ SO (3), avec le noyau que forment les matrices $\pm E$ (voir chap. 7, § 1), toute représentation Ψ du groupe SO (3) peut être considérée comme étant aussi une représentation du groupe SU (2) (voir démonstration du théorème 5 du § 5) qui satisfait à la *condition* dite *de parité*: $\Psi_{-E} = \Psi_E$. Ceci étant l'égalité $\Psi_{-g} = \Psi_g$ sera vérifiée bien entendu par tout $g \in$ $\in$ SU (2). Réciproquement, lorsque la condition de parité est satisfaite, la représentation Ψ du groupe SU (2) est en même temps une représentation du groupe SO (3). Des représentations « à deux signes » de SO (3), c'est-à-dire les représentations du groupe SU (2) qui ne satisfont pas à la condition de parité, ont, elles aussi, un sens physique. Parmi de telles représentations, signalons, par exemple, la représentation ordinaire de dimension deux (représentation spinorielle).

Remarquons encore que toute représentation irréductible du groupe SO (3), autre que la représentation unité, est exacte, ce qui résulte directement du fait que SO (3) est un groupe simple (chap. 7, § 3, théorème 6).

THÉORÈME 1. — *Soit* $V_n = \langle x^k y^{n-k} \mid k = 0, 1, \ldots, n \rangle_{\mathbf{C}}$ *l'espace des polynômes homogènes de degré n à deux indéterminées complexes, subissant une opération* $\Psi^{(n)}$ *du groupe* SU (2) *définie par la relation*

$$(\Psi_g^{(n)} f)(x, y) = f(\bar{\alpha}x - \beta y, \bar{\beta}x + \alpha y)$$

pour tout élément

$$g = \begin{Vmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{Vmatrix}, \quad |\alpha|^2 + |\beta|^2 = 1.$$

Alors $(\Psi^{(n)}, V_n)$ *est une représentation irréductible de dimension* $n+1$ *du groupe* SU (2). *Lorsque n est pair,* $(\Psi^{(n)}, V_n)$ *est aussi une représentation irréductible du groupe* SO (3).

DÉMONSTRATION. — Supposons que le polynôme

$$f(x, y) = \sum_{k=0}^{n} a_k x^k y^{n-k} \neq 0$$

appartienne à un sous-espace invariant $U \subset V_n$. Il vient alors que

$$\sum_{k=0}^{n} (e^{-i\varphi})^k a_k x^k y^{n-k} = e^{-in\frac{\varphi}{2}} (\Psi_{b_\varphi}^{(n)} f)(x, y) \in U,$$

où b_φ est un élément de SU (2) de la forme (4) (chap. 7, § 1). Etant donné que φ est un nombre réel arbitraire compris dans l'intervalle $(0, 2\pi)$, on peut composer un système linéaire, avec le déterminant de Vandermonde, dont il résulte que

$$f(x, y) \in U \Rightarrow x^k y^{n-k} \in U \qquad (2)$$

pour tout monôme à coefficient $a_k \neq 0$. Or, si $x^k y^{n-k} \in U$ pour un k quelconque, on a aussi

$$\bar{\alpha}^k \bar{\beta}^{n-k} x^m + \ldots = (\bar{\alpha}x - \beta y)^k (\bar{\beta}x + \alpha y)^{n-k} = \Psi_g^{(n)} (x^k y^{n-k}) \in U.$$

En prenant g, avec $\alpha\beta \neq 0$, nous arrivons en raison de (2) à l'appartenance $x^n \in U$ qui donne à son tour

$$\sum_{s=0}^{n} \binom{n}{s} \bar{\alpha}^s (-\beta)^{n-s} x^s y^{n-s} \in U.$$

Puisque $\binom{n}{s} \bar{\alpha}^s (-\beta)^{n-s} \neq 0$, on a $x^s y^{n-s} \in U$, $s = 0, 1, \ldots, n$. Par conséquent, $U = V_n$ et l'irréductibilité de $(\Psi^{(n)}, V_n)$ est démontrée.

On a ensuite

$$\Psi_{-E}^{(n)}(x^k y^{n-k}) = (-x)^k (-y)^{n-k} = (-1)^n x^k y^{n-k},$$

si bien que pour $n = 2m$, la condition de parité se trouve satisfaite (voir remarque ci-dessus) et la représentation $(\Psi^{(2m)}, V_{2m})$ du groupe SO (3) peut être considérée comme étant irréductible de dimension $2m + 1$. ■

Or, en réalité, $\Psi^{(2m)}$ est équivalente à la représentation $\Phi^{(m)}$ du groupe SO (3) sur l'espace des polynômes harmoniques homogènes de degré m, mais nous ne nous étendons pas sur ce sujet et n'essayons pas non plus de choisir dans V_n (bien que cela soit possible) une base telle que la représentation $\Psi^{(n)}$ devienne unitaire. Remarquons seulement, en utilisant la terminologie de l'Analyse tensorielle, que la représentation $\Psi^{(n)}$ du groupe SU (2) se réalise aussi dans la classe des tenseurs symétriques covariants de rang n. Une théorie complète et suffisamment transparente des représentations des groupes compacts, y compris SU (2) et SO (3), est généralement développée dans le cadre de la méthode infinitésimale qui s'appuie sur la correspondance entre les groupes et les algèbres de Lie.

EXERCICES

1. Construire $2m + 1$ polynômes harmoniques homogènes linéairement indépendants de degré m.

2. Montrer que tout polynôme homogène $f \in P_m$ s'écrit sous la forme d'une combinaison linéaire, à coefficients dépendant de $x^2 + y^2 + z^2$, de polynômes harmoniques de degrés m, $m - 2$, $m - 4$, ... (Indication. En comparant les dimensions, obtenir la décomposition de P_m en somme directe des espaces:

$$P_m = H_m \oplus (x^2 + y^2 + z^2) H_{m-2} \oplus (x^2 + y^2 + z^2)^2 H_{m-4} \oplus \ldots)$$

3. En partant de l'exercice 2, montrer que toute fonction polynomiale $\tilde{g}: (X, Y, Z) \mapsto g(x, y, z)$ sur la sphère $S^2: x^2 + y^2 + z^2 = 1$ se décompose suivant les *fonctions sphériques* qui sont des restrictions des polynômes harmoniques à S^2.

4. Montrer, sans avoir recours à la description complète des représentations irréductibles du groupe SO (3), que l'homomorphisme τ: SO (3) $\to$ SU (2) ne peut être que trivial. (Indication. Le groupe SO (3) étant simple, la non-trivialité de τ signifierait que τ est une représentation exacte de degré 2. Or, comme il résulte du § 5, n° 4, exemple 3, ou de la description des sous-groupes finis de SU (2) (voir § 3), même la restriction $\tau|_{S_4 \cong O}$ ne peut pas être exacte.)

§ 7. Produit tensoriel de représentations

1. Représentation duale.— Soit (Φ, V) une représentation d'un groupe G sur le corps $\mathbb{C}$. Introduisons l'espace dual V^* (espace des fonctions linéaires sur V) et posons

$$(\Phi^*(g) \cdot f)(v) = f(\Phi(g^{-1}) v); \quad f \in V^*, \quad g \in V. \qquad (1)$$

On vérifie immédiatement que l'opérateur $\Phi^*(g)$ est linéaire. Choisissons maintenant dans V et V^* des bases duales

$$V = \langle e_1, \ldots, e_n \rangle, \quad V^* = \langle f_1, \ldots, f_n \rangle, \quad f_i(e_j) = \delta_{ij}.$$

La matrice de l'opérateur linéaire $\Phi^*(g)$ dans la base $f_1, \ldots, f_n$ est la transposée de la matrice de l'opérateur $\Phi(g^{-1})$ dans la base $e_1, \ldots, e_n$:

$$\Phi_g^* = {}^t\Phi_{g^{-1}}. \tag{2}$$

Puisque

$$\Phi_{gh}^* = {}^t\Phi_{(gh)^{-1}} = {}^t\Phi_{h^{-1}g^{-1}} = {}^t(\Phi_{h^{-1}}\Phi_{g^{-1}}) = {}^t\Phi_{g^{-1}}{}^t\Phi_{h^{-1}} = \Phi_g^*\Phi_h^*,$$

la relation (2) (ou (1)) définit, en général, une nouvelle représentation linéaire (Φ^*, V^*) du groupe G; elle s'appelle représentation *duale* (ou *contragrédiente*) de (Φ, V). La nécessité de considérer de telles représentations apparaît chaque fois que nous faisons agir sur les coordonnées des vecteurs (tenseurs covariants) le groupe opérant sur les vecteurs (tenseurs contravariants), comme cela a été réellement fait au § 6. Il n'est pas difficile d'établir, ne serait-ce qu'à partir de (2), que $(\Phi^*)^* \approx \Phi$. Les représentations duales peuvent ne pas différer ou être équivalentes. Si par exemple (Φ, G) est une représentation orthogonale réelle, $\Phi_g^* = {}^t\Phi_{g^{-1}} = \Phi_g$. Mais, dans le cas général, les représentations Φ^* et Φ ne sont pas équivalentes, comme en témoigne un exemple bien simple

$$C_3 = \langle a \mid a^3 = e \rangle;\ \Phi(a) = \varepsilon,\ \Phi^*(a) = \varepsilon^{-1}\ (\varepsilon^2 + \varepsilon + 1 = 0).$$

Pour un groupe fini G, le critère rigoureux d'équivalence des représentations duales est obtenu en langage de théorie des caractères. Puisque les polynômes caractéristiques des matrices A et ${}^tA$ coïncident:

$$\det(\lambda E - {}^tA) = \det{}^t(\lambda E - A) = \det(\lambda E - A),$$

les propriétés élémentaires des caractères (proposition du § 4) entraînent que

$$\chi_{\Phi^*}(g) = \overline{\chi_\Phi(g)}.$$

En particulier, la représentation Φ dont le caractère ne prend que des valeurs réelles est équivalente à Φ^*. Bien entendu, l'égalité

$$(\chi_{\Phi^*}, \chi_{\Phi^*})_G = (\chi_\Phi, \chi_\Phi)_G$$

est toujours vérifiée, de sorte que Φ^* et Φ sont simultanément réductibles ou irréductibles.

2. Produit tensoriel de représentations.— Dans le cours d'Algèbre linéaire et de Géométrie (voir aussi l'exercice 1) on démontre l'assertion suivante:

THÉORÈME 1. — *Soient V, W deux espaces vectoriels sur un corps commutatif P. Alors, il existe un espace vectoriel T sur P et une application bilinéaire $\tau: V \times W \to T$ satisfaisant aux conditions suivantes:*

(T1) *si* $v_1, \ldots, v_k \in V$ *sont linéairement indépendants et* $w_1, \ldots$
$\ldots, w_k \in W$, *alors* $\sum_{i=1}^{k} \tau(v_i, w_i) = 0 \Rightarrow w_1 = 0, \ldots, w_k = 0$;

(T2) *si* $w_1, \ldots, w_k \in W$ *sont linéairement indépendants, alors* $\sum_i \tau(v_i, w_i) = 0 \Rightarrow v_1 = 0, \ldots, v_k = 0$;

(T3) $T = \langle \tau(v, w) \mid v \in V,\ w \in W \rangle_P$.

En outre, le couple (τ, T) *est universel en ce sens que, quel que soit le couple* (τ', T') *constitué de l'espace vectoriel* T' *et de l'application bilinéaire* $\tau': V \times W \to T'$, *il existe une application linéaire et une seule* $\sigma: T \to T'$ *telle que*

$$\tau'(v, w) = \sigma(\tau(v, w)), \quad v \in V, \quad w \in W. \blacksquare$$

En supposant l'existence de deux couples universels (τ, T), (τ', T'), nous constatons sans peine qu'en réalité les applications linéaires $\sigma: T \to T'$, $\sigma': T' \to T$ sont des isomorphismes réciproques: $\sigma' \circ \sigma = e_T$, $\sigma \circ \sigma' = e_{T'}$. Ainsi, $T \cong T'$ et l'isomorphisme $\sigma: T \to T'$ vérifie la propriété indiquée dans l'énoncé du théorème.

Le couple (τ, T), défini de façon unique, à un isomorphisme près, par les espaces vectoriels donnés V, W, est appelé *produit tensoriel* de ces espaces. En notant $T = V \otimes_P W$ ou tout simplement $T = V \otimes W$, on ne devra pas oublier que l'espace vectoriel T est muni d'une application bilinéaire $(v, w) \mapsto v \otimes w$ du produit cartésien $V \otimes W$ dans T, qui possède les propriétés (T1) à (T3). Ainsi, les éléments du produit tensoriel $V \otimes W$ sont des combinaisons linéaires formelles, à coefficients dans P, des couples ordonnés $v \otimes w$, avec $v \in V$, $w \in W$. Ceci étant, on suppose que les conditions suivantes sont vérifiées:

$$\begin{gathered}(v_1 + v_2) \otimes w - v_1 \otimes w - v_2 \otimes w = 0, \\ v \otimes (w_1 + w_2) - v \otimes w_1 - v \otimes w_2 = 0, \\ \lambda(v \otimes w) - \lambda v \otimes w = 0 = \lambda v \otimes w - v \otimes \lambda w, \quad \lambda \in P.\end{gathered} \tag{3}$$

Du théorème 1 on déduit immédiatement que les applications bijectives $v \otimes w \mapsto w \otimes v$, $(u \otimes v) \otimes w \mapsto u \otimes (v \otimes w)$, $v \otimes \lambda \mapsto \lambda \otimes v \mapsto \lambda v$ établissent des isomorphismes dits *canoniques* des espaces vectoriels:

$$\begin{gathered}V \otimes W \cong W \otimes V, \\ (U \otimes V) \otimes W \cong U \otimes (V \otimes W), \\ V \otimes P \cong P \otimes V \cong V.\end{gathered}$$

Les lois de distributivité sont, elles aussi, vérifiées:

$$\begin{gathered}(U \oplus V) \otimes W \cong (U \otimes W) \oplus (V \otimes W), \\ U \otimes (V \oplus W) \cong (U \otimes V) \oplus (U \otimes W).\end{gathered}$$

En Analyse tensorielle, d'où tiennent leur origine les notions que nous sommes en train de considérer, on étudie des produits tensoriels de forme spéciale

$$\underbrace{V^* \otimes \ldots \otimes V^*}_{p} \otimes \underbrace{V \otimes \ldots \otimes V}_{q}.$$

Leurs éléments sont des tenseurs de type (p, q), p fois covariants et q fois contravariants. Dans le cas où l'on choisit les bases duales $e_1, \ldots, e_n$ de V et $e^1, \ldots, e^n$ de V^*, les éléments $e^{i_1} \otimes \ldots \otimes e^{i_p} \otimes e_{j_1} \otimes \ldots \otimes e_{j_q}$ constituent une base de l'espace des tenseurs de type (p, q). Généralement, on entend par tenseur tout simplement un ensemble de coordonnées $\{t_{i_1 \ldots i_p}^{j_1 \ldots j_q}\}$ dans cette base, en indiquant les règles de changement de coordonnées lors du passage d'une base à l'autre. C'est ainsi que l'on interprète en langage tensoriel (au fait, en langage de matrices) des notions telles que forme bilinéaire et opérateur linéaire. En nous bornant à ces brefs rappels, que nous n'avons pas l'intention d'utiliser en pleine mesure, revenons à la question de représentations qui nous intéresse.

Soient $\mathcal{A}: V \to V$, $\mathcal{B}: W \to W$ deux opérateurs linéaires. On appelle *produit tensoriel* de $\mathcal{A}$ et $\mathcal{B}$ l'opérateur linéaire

$$\mathcal{A} \otimes \mathcal{B}: V \otimes W \to V \otimes W,$$

agissant suivant la loi

$$(\mathcal{A} \otimes \mathcal{B})(v \otimes w) = \mathcal{A}v \otimes \mathcal{B}w \tag{4}$$

(puis par linéarité: $(\mathcal{A} \otimes \mathcal{B})(\sum v_i \otimes w_i) = \sum \mathcal{A}v_i \otimes \mathcal{B}w_i$). Il est clair que cette définition est en accord avec les relations (3). Par exemple,

$$\mathcal{A}(v_1 + v_2) \otimes \mathcal{B}w - \mathcal{A}v_1 \otimes \mathcal{B}w - \mathcal{A}v_2 \otimes \mathcal{B}w =$$
$$= (\mathcal{A}v_1 + \mathcal{A}v_2) \otimes \mathcal{B}w - \mathcal{A}v_1 \otimes \mathcal{B}w - \mathcal{A}v_2 \otimes \mathcal{B}w = 0.$$

Aussi, l'opération $\mathcal{A} \otimes \mathcal{B}$ sur $V \otimes W$ est-elle définie correctement. Signalons aussi les relations suivantes qui résultent immédiatement de la définition (4):

$$(\mathcal{A} \otimes \mathcal{B})(\mathcal{C} \otimes \mathcal{D}) = \mathcal{A}\mathcal{C} \otimes \mathcal{B}\mathcal{D},$$
$$(\mathcal{A} + \mathcal{C}) \otimes \mathcal{B} = \mathcal{A} \otimes \mathcal{B} + \mathcal{C} \otimes \mathcal{B},$$
$$\mathcal{A} \otimes (\mathcal{B} + \mathcal{D}) = \mathcal{A} \otimes \mathcal{B} + \mathcal{A} \otimes \mathcal{D},$$
$$\mathcal{A} \otimes \lambda\mathcal{B} = \lambda\mathcal{A} \otimes \mathcal{B} = \lambda(\mathcal{A} \otimes \mathcal{B}).$$

Nous laissons au lecteur le soin de les vérifier.

Soit, comme précédemment, $V = \langle e_1, \ldots, e_n \rangle$, $W = \langle f_1, \ldots, f_m \rangle$. Quant à la matrice $A \otimes B$ de type $nm \times nm$ de l'opé-

rateur $\mathcal{A} \otimes \mathcal{B}$ relative à la base

$$\{e_1 \otimes f_1, \ldots, e_1 \otimes f_m, e_2 \otimes f_1, \ldots, e_2 \otimes f_m, \ldots \\ \ldots, e_n \otimes f_1, \ldots, e_n \otimes f_m\},$$

nous obtenons en remarquant que

$$\mathcal{A}e_i = \sum_{i'} \alpha_{i'i} e_{i'}, \quad \mathcal{B}f_j = \sum_{j'} \beta_{j'j} f_{j'},$$

$$(\mathcal{A} \otimes \mathcal{B})(e_i \otimes f_j) = \sum_{i', j'} \alpha_{i'i}\beta_{j'j} e_{i'} \otimes f_{j'}.$$

Ainsi donc, nous avons

$$A \otimes B = (\alpha_{i'i}\beta_{j'j}) = \begin{Vmatrix} \alpha_{11}B & \alpha_{12}B & \ldots & \alpha_{1n}B \\ \alpha_{21}B & \alpha_{22}B & \ldots & \alpha_{2n}B \\ \ldots & \ldots & \ldots & \ldots \\ \alpha_{n1}B & \alpha_{n2}B & \ldots & \alpha_{nn}B \end{Vmatrix},$$

avec $A = (\alpha_{i'i})$, $B = (\beta_{j'j})$. En particulier, nous avons pour la trace la formule suivante:

$$\operatorname{tr} A \otimes B = \alpha_{11} \operatorname{tr} B + \alpha_{22} \operatorname{tr} B + \ldots + \alpha_{nn} \operatorname{tr} B = \\ = \operatorname{tr} A \cdot \operatorname{tr} B. \qquad (5)$$

Remarquons, en passant, que

$$\det A \otimes B = \det (A \otimes E_m)(E_n \otimes B) = \\ = \det A \otimes E_m \cdot \det E_n \otimes B = (\det A)^m (\det B)^n,$$

de sorte que, si les opérateurs $\mathcal{A}$ et $\mathcal{B}$ ne sont pas dégénérés, il en est de même de leur produit tensoriel.

Soient maintenant (Φ, V), (Ψ, W) deux représentations linéaires d'un groupe G de caractères χ_Φ et χ_Ψ respectivement. Définissons de façon naturelle la représentation $(\Phi \otimes \Psi, V \otimes W)$, en posant

$$(\Phi \otimes \Psi)(g) = \Phi(g) \otimes \Psi(g), \quad \forall g \in G.$$

Les propriétés générales du produit tensoriel des opérateurs linéaires, jointes à la formule (5), entraînent que l'application $\Phi \otimes \Psi$ définit, en effet, une représentation du groupe G avec l'espace de représentation $V \otimes W$ et le caractère

$$\chi_{\Phi \otimes \Psi} = \chi_\Phi \chi_\Psi. \qquad (6)$$

Nous dirons que $(\Phi \otimes \Psi, V \otimes W)$ est *produit tensoriel des représentations* (Φ, V) et (Ψ, W). Pour $\Psi = \Phi$, $W = V$ on dit aussi que l'on a affaire à un *carré tensoriel*. Le second membre de la formule (6) est le produit ordinaire des fonctions centrales χ_Φ et χ_Ψ.

Il est tout à fait évident que, si U est un sous-espace G-invariant de V, alors $U \otimes W$ sera un sous-espace G-invariant de $V \otimes W$. Une remarque analogue est valable pour les sous-espaces G-invariants

de W. Cependant, comme le montre l'exemple du carré tensoriel $\Phi^{(3)} \otimes \Phi^{(3)}$ de la représentation de dimension deux du groupe S_3 (voir table au § 5, nº 2), l'irréductibilité de V et W n'entraîne aucunement celle de $V \otimes W$. En effet, $\dim_{\mathbb{C}} \Phi^{(3)} \otimes \Phi^{(3)} = 4$, tandis que le degré maximal de la représentation irréductible du groupe S_3 est égal à 2.

Le problème de description efficace des représentations irréductibles contenues dans $\Phi \otimes \Psi$ et, plus généralement, dans le produit tensoriel $\Phi^{(1)} \otimes \Phi^{(2)} \otimes \ldots \otimes \Phi^{(p)}$ de plusieurs représentations linéaires, revêt une importance de principe, car de nombreuses représentations importantes et bien naturelles des groupes apparaissent comme produits tensoriels. C'est justement de ce point de vue qu'il faut considérer les représentations des groupes SU (2) et SO (3) (voir § 6), ainsi que les exemples 3 et 4 du § 1, nº 2. Les sous-espaces invariants des tenseurs covariants (ou contravariants) symétriques et symétriques gauches se rencontrent constamment dans les diverses applications géométriques. Le problème considéré est surtout attrayant dans le cas où le théorème sur la réductibilité complète des représentations est valable.

3. Anneau de caractères.— Par raison de simplification, bornons-nous à examiner le cas d'un groupe fini G et du corps $\mathbb{C}$. Soient $\Phi^{(1)}, \Phi^{(2)}, \ldots, \Phi^{(r)}$ l'ensemble complet des représentations irréductibles deux à deux non équivalentes du groupe G sur $\mathbb{C}$, et $\chi_1, \chi_2, \ldots, \chi_r$ les caractères correspondants de ces représentations (r est le nombre de classes d'éléments conjugués dans G). Nous savons que

$$\Phi \otimes \Psi \approx m_1\Phi^{(1)} \dot{+} \ldots \dot{+} m_r\Phi^{(r)},$$

où les ordres de multiplicité m_i ne dépendent que de Φ et Ψ. D'après la formule (6) on a :

$$\chi_\Phi \chi_\Psi = m_1\chi_1 + \ldots + m_r\chi_r.$$

Soit $X_{\mathbb{Z}}(G)$ l'ensemble de toutes les combinaisons linéaires possibles, à coefficients entiers, des caractères $\chi_1, \ldots, \chi_r$. Nous avons démontré plus haut que $\chi_1, \ldots, \chi_r$ est une base orthonormée de l'espace $X_{\mathbb{C}}(G)$, et de ce fait $X_{\mathbb{Z}}(G) \subset X_{\mathbb{C}}(G)$ est en tout cas un groupe abélien libre $\cong \mathbb{Z}^r$ engendré par $\chi_1, \ldots, \chi_r$. Ses éléments s'appellent *caractères généralisés* du groupe G. Les seuls vrais caractères seront les combinaisons $\sum m_i\chi_i$, avec $m_i \geqslant 0$.

Il résulte de tout ce qui précède que le produit tensoriel des représentations induit sur $X_{\mathbb{Z}}(G)$ une opération algébrique binaire qui est commutative, associative et distributive par rapport à l'addition. Bref, on a le théorème suivant :

THÉORÈME 2.— *Les caractères généralisés forment un anneau commutatif et associatif $X_{\mathbb{Z}}(G)$ ayant un élément unité qui est caractère unité χ_1.* ■

$X_{\mathbb{C}}(G)$ s'appelle algèbre associative et commutative de dimension r sur $\mathbb{C}$. La structure de l'anneau $X_{\mathbb{Z}}(G)$ (de l'algèbre $X_{\mathbb{C}}(G)$) est entièrement déterminée par les *constantes de structure*, c'est-à-dire par les entiers m_{ij}^k figurant dans les relations

$$\chi_i\chi_j = \sum m_{ij}^k \chi_k. \tag{7}$$

En particulier, les égalités $m_{ij}^k = m_{ji}^k$, $m_{1j}^k = \delta_{kj}$ reflètent la propriété de commutativité de $X_{\mathbb{Z}}(G)$ et le fait que χ_1 est l'élément unité. D'après (7), on a

$$\chi_i(g)\chi_j(g) = \sum m_{ij}^k \chi_k(g), \quad \forall g \in G.$$

Multiplions les deux membres de cette relation par $\frac{1}{|G|}\overline{\chi_s(g)}$ et effectuons la sommation sur $g \in G$. En utilisant la première relation d'orthogonalité pour les caractères, nous obtenons

$$m_{ij}^s = \frac{1}{|G|}\sum_{g\in G} \chi_i(g)\chi_j(g)\overline{\chi_s(g)}. \tag{8}$$

Ainsi, les constantes de structure s'expriment en termes des caractères eux-mêmes.

De (8) on peut tirer une assertion peu compliquée, à savoir:

$$m_{ij}^1 = \frac{1}{|G|}\sum_g \chi_i(g)\chi_j(g)\overline{\chi_1(g)} = \frac{1}{|G|}\sum_g \chi_i(g)\chi_j(g) =$$
$$= \frac{1}{|G|}\sum_g \chi_i(g)\overline{\overline{\chi_j(g)}} = (\chi_i, \chi_j^*)_G,$$

où $\chi_j^* = \chi_{\Phi^{(j)*}}$ est le caractère de la représentation duale de $\Phi^{(j)}$ (voir nº 1). Ainsi, la *représentation unité apparaît comme composante dans la décomposition de* $\Phi^{(i)} \otimes \Phi^{(j)}$ *si, et seulement si,* $\Phi^{(i)}$ *est équivalente à la représentation* $\Phi^{(j')} = \Phi^{(j)*}$ (dans le cas contraire, $m_{ij}^1 = (\chi_i, \chi_j^*)_G = 0$). ■

Signalons encore que le *produit tensoriel d'une représentation de dimension un* $\Phi^{(i)}$ *et d'une représentation irréductible arbitraire* $\Phi^{(j)}$ *est toujours une représentation irréductible de même dimension que* $\Phi^{(j)}$. Cette assertion suffisamment évidente résulte formellement du critère d'irréductibilité des caractères. Si $\chi = \chi_{\Phi^{(i)}\otimes\Phi^{(j)}} = \chi_i\chi_j$, alors $\chi_i(g)$ est une racine complexe d'une certaine puis-

sance de l'unité, et $\chi_i(g)\overline{\chi_i(g)} = 1$, de sorte que

$$(\chi, \chi)_G = \frac{1}{|G|}\sum_g \chi_i(g)\chi_j(g)\overline{\chi_i(g)}\,\overline{\chi_j(g)} =$$
$$= \frac{1}{|G|}\sum_g \chi_j(g)\overline{\chi_j(g)} = (\chi_j, \chi_j)_G = 1. \quad \blacksquare$$

EXEMPLE 1. $G = S_3$ (voir tables au § 2, n° 1 et au § 5, n° 2):

$$\Phi^{(1)} \otimes \Phi^{(3)} \approx \Phi^{(2)} \otimes \Phi^{(3)} \approx \Phi^{(3)}.$$

EXEMPLE 2. $G = S_4$ (voir § 5, n° 4, exemple 3):

$$\Phi^{(2)} \otimes \Phi^{(4)} \approx \Phi^{(5)}, \quad \Phi^{(2)} \otimes \Phi^{(5)} \approx \Phi^{(4)}.$$

Enfin, démontrons un théorème assez intéressant qui généralise le théorème 2 du § 5 sur la décomposition d'une représentation régulière.

THÉORÈME 3. — *Soit $\chi = \chi_\Phi$ le caractère de la représentation exacte (Φ, V) d'un groupe fini G sur le corps des nombres complexes $\mathbb{C}$ prenant sur G exactement m valeurs différentes. Alors, chaque caractère irréductible χ_k apparaît avec un coefficient non nul dans la décomposition d'au moins un caractère $\chi^0 = \chi_1, \chi, \chi^2, \ldots, \chi^{m-1}$. En d'autres termes, toute représentation irréductible est contenue dans la décomposition d'une certaine puissance tensorielle $\Phi^{\otimes i} = \Phi \otimes \ldots \otimes \Phi$, $0 \leqslant i \leqslant m-1$, de toute représentation exacte Φ.*

DÉMONSTRATION. — Soient $\omega_j = \chi(g_j)$, $j = 0, 1, \ldots, m-1$, les différentes valeurs prises par le caractère χ sur G, avec $\omega_0 = \chi(e) = \deg \Phi$. Soit ensuite

$$G_j = \{g \in G \mid \chi(g) = \chi(g_j) = \omega_j\}.$$

Du fait que la représentation Φ est exacte, on a

$$G_0 = \operatorname{Ker} \Phi = \{e\}.$$

Soit χ_k un caractère irréductible du groupe G qui n'intervient dans la décomposition d'aucun des caractères χ^i. Alors

$$0 = |G|\,(\chi^i, \chi_k)_G = \sum_{j=0}^{m-1} (\chi(g_j))^i \sum_{g \in G_j} \overline{\chi_k(g)} = \sum \omega_j^i T_j, \quad 0 \leqslant i \leqslant m-1,$$

est un système homogène d'équations linéaires par rapport à $T_j = \sum_{g \in G_j} \overline{\chi_k(g)}$, avec le déterminant

$$\det(\omega_j^i) = \begin{vmatrix} 1 & 1 & \ldots & 1 \\ \omega_0 & \omega_1 & \ldots & \omega_{m-1} \\ \cdots & \cdots & \cdots & \cdots \\ \omega_0^{m-1} & \omega_1^{m-1} & \ldots & \omega_{m-1}^{m-1} \end{vmatrix}$$

différent de zéro (déterminant de Vandermonde). Ainsi, $T_j = 0$, $j = 0, 1, \ldots, m-1$, c'est-à-dire

$$\sum_{g \in G_j} \chi_k(g^{-1}) = 0, \qquad j = 0, 1, \ldots, m-1.$$

En particulier,

$$0 = \sum_{g \in G_0} \chi_k(g^{-1}) = \chi_k(e),$$

contradiction qui démontre le théorème. ■

Dans le cas d'une représentation régulière ρ, on a évidemment $m = 2$.

4. Invariants des groupes linéaires.— Comme à l'habitude, nous appelons groupe linéaire de degré n tout sous-groupe de $GL(n, K)$, où K est un corps commutatif. Dans la suite, on peut admettre que $K = \mathbb{R}$ ou $\mathbb{C}$. Si G est un groupe abstrait et $\Phi: G \to GL(n, \mathbb{C})$ est sa représentation linéaire, le couple (G, Φ) sera appele, lui aussi, groupe linéaire. Les transformations linéaires Φ_g opèrent sur les colonnes des variables $x_1, \ldots, x_n$, ce qu'on représente conventionnellement sous la forme suivante:

$$\begin{Vmatrix} \Phi_g(x_1) \\ \vdots \\ \Phi_g(x_n) \end{Vmatrix} = \Phi_g \begin{Vmatrix} x_1 \\ \vdots \\ x_n \end{Vmatrix}.$$

L'image par Φ_g de toute forme (polynôme homogène) f de degré m est encore une forme de degré m:

$$(\tilde{\Phi}_g f)(x_1, \ldots, x_n) = f(\Phi_{g^{-1}}(x_1), \ldots, \Phi_{g^{-1}}(x_n)).$$

Nous avons déjà rencontré (voir § 6) divers cas particuliers de cette opération. L'application $\tilde{\Phi}$ détermine la représentation du groupe G sur l'espace P_m des formes sur $\mathbb{C}$ de degré m (ou des tenseurs symétriques covariants de rang m).

Définition. — *Une forme $f \in P_m$ qui reste stable par $\tilde{\Phi}_g$ ($\tilde{\Phi}_g f = f$, $\forall g \in G$) s'appelle invariant (entier) de degré m du groupe linéaire (G, Φ).*

Strictement parlant, il faudrait prendre un polynôme de coefficients de la forme « générale » de degré m, qui reste stable lors de l'opération $\tilde{\Phi}(G)$. C'est ce qu'on fait en théorie générale des invariants, mais par souci de simplification nous nous contenterons de la définition donnée. Si l'on prend pour f une fonction rationnelle, on peut arriver à la notion d'*invariant rationnel*. On utilise aussi la notion importante d'*invariant relatif* f, lorsque

$$\tilde{\Phi}_g f = \omega_g f,$$

où $\omega_g \in \mathbb{C}$ est un facteur dépendant de l'élément $g \in G$.

Il est clair que tout ensemble $\{f_1, f_2, \ldots\}$ des invariants du groupe (G, Φ) engendre dans $\mathbb{C}[x_1, \ldots, x_n]$ un sous-anneau $\mathbb{C}[f_1, f_2, \ldots]$ des invariants.

Considérons quelques exemples.

EXEMPLE 1.— La forme quadratique $x_1^2 + x_2^2 + \ldots + x_n^2$ et tous les polynômes en fonction de cette forme sont des invariants entiers du groupe orthogonal $O(n)$.

EXEMPLE 2. — Les polynômes symétriques élémentaires $s_1(x_1, \ldots, x_n), \ldots, s_n(x_1, \ldots, x_n)$ sont des invariants entiers du groupe symétrique S_n considéré avec le monomorphisme canonique $\Phi: S_n \to GL(n)$. Le théorème fondamental sur les polynômes symétriques dit que les invariants $s_1, \ldots, s_n$ de degrés $1, 2, \ldots, n$ sont algébriquement indépendants et que les fonctions polynomiales (rationnelles) de ces invariants sont les seuls invariants entiers (rationnels) du groupe (S_n, Φ).

Les invariants relatifs du groupe linéaire (S_n, Φ) sont représentés par les polynômes antisymétriques: $\Phi_\pi f = (\det \Phi_\pi) f = \varepsilon_\pi f$. Nous avons vu (chap. 6, § 2, exercice 3) que tout polynôme antisymétrique f est de la forme $f = \Delta_n \cdot g$, où $\Delta_n = \prod_{j<i} (x_i - x_j)$ et g est un polynôme symétrique arbitraire, c'est-à-dire un invariant absolu.

EXEMPLE 3.— A la représentation $\Phi_A: X \to AXA^{-1}$ de degré n^2 du groupe linéaire complet $GL(n, K)$, avec l'espace de représentation $M_n(K)$ (voir § 1, exemple 3), correspond un système de n invariants qui sont algébriquement indépendants, c'est-à-dire le système des coefficients du polynôme caractéristique de la matrice $X = (x_{ij})$. Parmi ces invariants se rangent en particulier les invariants bien connus: $\operatorname{tr} X = \sum x_{ii}$ et $\det X$.

EXEMPLE 4.— Le groupe orthogonal $O(n)$ opère sur la forme quadratique $f(x_1, \ldots, x_n) = \sum a_{ij} x_i x_j$, écrite sous la forme $f(x_1, \ldots, x_n) = {}^tXAX$, $A = (a_{ij}) = {}^tA$, $X = [x_1, \ldots, x_n]$:

$$C \in O(n) \Rightarrow (C^{-1}f)(x_1, \ldots, x_n) = {}^t(CX)A(CX) = {}^tX{}^tCACX = {}^tX(C^{-1}AC)X.$$

Dans ce cas on convient de dire que l'on a affaire aux invariants de la forme quadratique f par rapport à $O(n)$: $\operatorname{tr} A, \ldots, \det A$. Pour la forme quadratique binaire $ax^2 + 2bxy + cy^2$, les invariants $a + c$ et $ac - b^2$, caractérisant les classes métriquement différentes de courbes du deuxième ordre, sont connus encore du cours de Géométrie analytique.

EXEMPLE 5.— Considérons le groupe symétrique S_3 comme groupe linéaire de degré 2, en utilisant une représentation Γ équivalente à la représentation $\Phi^{(3)}$ indiquée dans la table donnée à la fin du n° 1 du § 2:

$$\Gamma_{(123)} = \begin{Vmatrix} \varepsilon & 0 \\ 0 & \varepsilon^{-1} \end{Vmatrix}, \quad \Gamma_{(23)} = \begin{Vmatrix} 0 & 1 \\ 1 & 0 \end{Vmatrix}, \quad \varepsilon^2 + \varepsilon + 1 = 0$$

(l'équivalence est assurée par l'intermédiaire de la conjugaison:

$$\begin{Vmatrix} \varepsilon & 0 \\ 0 & 1 \end{Vmatrix} \Phi_\sigma^{(3)} \begin{Vmatrix} \varepsilon^{-1} & 0 \\ 0 & 1 \end{Vmatrix} = \Gamma_\sigma \Big).$$

Soient u, v des variables indépendantes, linéairement transformées par Γ_σ:

$$\Gamma_{(123)}(u) = \varepsilon u, \ \Gamma_{(123)}(v) = \varepsilon^{-1} v; \ \Gamma_{(23)}(u) = v, \ \Gamma_{(23)}(v) = u.$$

Puisque

$$\widetilde{\Gamma}_{(123)}(uv) = \Gamma^{-1}_{(123)}(u)\,\Gamma^{-1}_{(123)}(v) = \varepsilon^{-1}u\cdot\varepsilon v = uv,$$
$$\widetilde{\Gamma}_{(23)}(uv) = vu = uv,$$
$$\widetilde{\Gamma}_{(123)}(u^3+v^3) = (\varepsilon^{-1}u)^3 + (\varepsilon v)^3 = u^3+v^3,$$
$$\widetilde{\Gamma}_{(23)}(u^3+v^3) = v^3+u^3 = u^3+v^3,$$

le groupe (S_3, Γ) possède les formes

$$I_1 = uv, \qquad I_2 = u^3+v^3 \tag{9}$$

de degrés 2 et 3 comme ses invariants.

Le groupe S_3 opère de façon naturelle sur les polynômes $f(x_1, x_2, x_3)$ à trois indéterminées indépendantes:

$$(\sigma f)(x_1, x_2, x_3) = f(x_{\sigma^{-1}(1)}, x_{\sigma^{-1}(2)}, x_{\sigma^{-1}(3)}).$$

En posant

$$u = x_1 + \varepsilon x_2 + \varepsilon^2 x_3, \quad v = x_1 + \varepsilon^2 x_2 + \varepsilon x_3, \tag{10}$$

nous voyons que

$$\Gamma_\sigma(u) = x_{\sigma^{-1}(1)} + \varepsilon x_{\sigma^{-1}(2)} + \varepsilon^2 x_{\sigma^{-1}(3)}.$$

En particulier,

$$\Gamma_{(123)}(u) = x_3 + \varepsilon x_1 + \varepsilon^2 x_2 = \varepsilon u, \qquad \Gamma_{(23)}(u) = x_1 + \varepsilon x_3 + \varepsilon^2 x_2 = v,$$
$$\Gamma_{(123)}(v) = x_3 + \varepsilon^2 x_1 + \varepsilon x_1 = \varepsilon^{-1} v, \qquad \Gamma_{(23)}(v) = x_1 + \varepsilon^2 x_3 + \varepsilon x_2 = u,$$

c'est-à-dire que les opérations de Γ_σ sur u, v et de σ sur x_1, x_2, x_3 sont compatibles. L'introduction de (10) dans les invariants (9) transforme ces derniers en fonctions symétriques de x_1, x_2, x_3, qui peuvent être exprimées, d'après le théorème 1 du chapitre 6, § 2, par les fonctions symétriques élémentaires $s_i = s_i(x_1, x_2, x_3)$.

Un petit exercice montre que

$$I_1 = x_1^2 + x_2^2 + x_3^2 + (\varepsilon + \varepsilon^2)(x_1x_2 + x_1x_3 + x_2x_3) = s_1^2 - 3s_2,$$
$$I_2 = 2(x_1^3 + x_2^3 + x_3^3) -$$
$$-3(x_1^2x_2 + x_1^2x_3 + x_1x_2^2 + x_1x_3^2 + x_2^2x_3 + x_2x_3^2) + 12x_1x_2x_3 = 2s_1^3 - 9s_1s_2 +$$
$$+ 27s_3.$$

Spécialisons les valeurs de I_1, I_2 en prenant pour x_1, x_2, x_3 trois racines de l'équation cubique incomplète

$$x^3 + px + q = 0.$$

Alors $s_1 = 0$, $s_2 = p$, $s_3 = -q$ et donc

$$I_1 = -3p, \qquad I_2 = -27q. \tag{11}$$

Or, il résulte de (9) que

$$v = \frac{I_1}{u}, \quad I_2 = u^3 + \frac{I_1^3}{u^3}, \quad u = \sqrt[3]{\frac{I_2}{2} \pm \sqrt{\frac{I_2^2}{4} - I_1}}$$

Tous les radicaux sont choisis de manière à obtenir, après l'introduction des valeurs (11), les formules

$$u = \sqrt[3]{-\frac{27}{2}q + \frac{3}{2}\sqrt{-3D}},$$
$$v = \sqrt[3]{-\frac{27}{2}q - \frac{3}{2}\sqrt{-3D}}, \quad uv = -3p,$$

avec $D = -4p^3 - 27q^2$ qui est le discriminant de notre équation cubique (voir chap. 6, § 2, relation (16)). Puisque u et v sont maintenant connues, le système linéaire

$$x_1 + \varepsilon x_2 + \varepsilon^2 x_3 = u,$$
$$x_1 + \varepsilon^2 x_2 + \varepsilon x_3 = v,$$
$$x_1 + x_2 + x_3 = 0$$

permet de trouver les racines elles-mêmes. Nous sommes arrivés d'une façon assez naturelle aux formules de Cardan qui ont été mentionnées au chapitre 1, § 2, problème 1.

Ce n'est pas un effet du hasard que le dernier exemple établit un lien entre les invariants du groupe S_3 qui est un groupe de Galois de l'équation cubique générale, et les formules de Cardan. La théorie de Galois est liée pour une large part à l'étude des invariants des corps commutatifs (et des groupes qui leur correspondent) engendrés par des racines des équations algébriques.

Signalons certains faits ayant trait au système de générateurs de l'anneau des invariants. Soit w une forme quelconque de n variables indépendantes $x_1, \ldots, x_n$. Le groupe fini G avec une représentation linéaire Φ de degré n opère comme un groupe de permutations sur l'ensemble

$$\Omega = \{\widetilde{\Phi}_g(w) \mid g \in G\}.$$

Il est clair que toute fonction symétrique homogène de $|G|$ (ou peut-être d'un certain diviseur du nombre $|G|$) éléments de Ω sera un invariant du groupe linéaire (G, Φ). Si l'on prend maintenant pour w la variable x_i, alors x_i sera racine de l'équation algébrique

$$\prod_{g \in G} (X - \Phi_g(x_i)) = 0$$

dont les coefficients sont des invariants du groupe (G, Φ). Ainsi, toute variable x_i est une fonction (algébrique) des invariants. Si le nombre d'invariants algébriquement indépendants était inférieur à n, nous aurions exprimé $x_1, \ldots, x_n$ par un plus petit nombre de variables algébriquement indépendantes, ce qui est impossible. Par conséquent, nous avons démontré (si l'on peut qualifier de « démonstration » un maniement si hardi de la dépendance algébrique des grandeurs) l'un des théorèmes importants de la théorie des invariants.

THÉORÈME 4. — *Un groupe linéaire fini de degré n possède toujours un système de n invariants algébriquement indépendants.* ■

Pour le groupe (S_3, Γ), ces invariants sont constitués par les formes (9).

On pourrait compléter le théorème 4 de l'assertion que tout l'anneau des invariants entiers d'un groupe linéaire fini de degré n est engendré par n invariants algébriquement indépendants

$f_1, f_2, \ldots, f_n$ et, en règle générale, encore par un invariant f_{n+1} (qui est une fonction algébrique des n premiers invariants). Autrement dit, tous les autres invariants sont des polynômes de $f_1, \ldots$ $\ldots, f_n, f_{n+1}$. Ce fait est vrai pour de nombreux autres groupes linéaires tant discrets que continus.

La théorie générale des invariants qui a été développée au milieu du XIXe siècle dans les ouvrages de Cayley, Sylvester, Jacobi, Hermite, Clebsch, Gordan et d'autres savants, et a eu ensuite une deuxième naissance dans quelques travaux fondamentaux de D. Hilbert, est devenue de nos jours une partie intégrante de la géométrie algébrique et de la théorie des groupes algébriques. Un intérêt permanent dont jouit la théorie des invariants s'explique aussi par de larges possibilités de ses applications dans de nombreuses branches de la mécanique et de la physique.

EXERCICES

1. Démontrer le théorème 1, en suivant les désignations utilisées dans son énoncé et les raisonnements schématisés ci-dessous.

a) Si $V = \langle e_1, \ldots, e_n \rangle_P$, $W = \langle f_1, \ldots, f_m \rangle_P$, alors (T1) à (T3) sont équivalentes dans leur ensemble à l'unique condition : les vecteurs $\tau(e_i, f_j)$, $1 \leqslant i \leqslant n$ $1 \leqslant j \leqslant m$, constituent une base de l'espace T.

b) Pour tout espace T de dimension nm sur P, l'application τ peut être définie par la relation $\tau(v, w) = \sum \alpha_i \beta_j g_{ij}$, où g_{ij}, $1 \leqslant i \leqslant n$, $i \leqslant j \leqslant m$, est une base de T. D'après a) le couple (τ, T) satisfait aux conditions (T1) à (T3), et tous les couples sont obtenus de cette manière.

c) Pour tout couple (τ', T'), avec $\tau' : V \times W \to T'$ une application bilinéaire, définissons une application linéaire $\sigma : T \to T'$, en posant $\sigma(\sum \gamma_{ij} g_{ij}) =$ $= \sum \gamma_{ij} \tau'(e_i, f_j)$.

D'après b) et c) on a $\tau'(v, w) = \sum \alpha_i \beta_j \tau'(e_i, f_j) = \sigma(\sum \alpha_i \beta_j g_{ij}) =$ $= \sigma(\tau(v, w))$. Réciproquement, si $\sigma(\tau(v, w)) = \tau'(v, w)$, alors $\sigma(g_{ij}) =$ $= \sigma(\tau(e_i, f_j)) = \tau'(e_i, f_j)$.

2. Montrer que dans le système (T1) à (T3) on peut omettre la condition (T1) ou (T2), et qu'en supposant *a priori* que $\dim T = nm$, il suffit, pour déterminer le produit tensoriel, de laisser l'une des trois conditions.

3. Démontrer la relation $\det(A \otimes B) = (\det A)^m \det(B)^n$ pour des matrices carrées A, B d'ordres respectifs n et m à coefficients complexes, en utilisant le fait qu'elles peuvent être réduites à la forme triangulaire (Indication. Il existe des matrices non singulières C et D telles que

$$A' = CAC^{-1} = \begin{Vmatrix} \alpha_1 & & * \\ & \ddots & \\ 0 & & \alpha_n \end{Vmatrix}, \quad B' = DBD^{-1} \begin{Vmatrix} \beta_1 & & * \\ & \ddots & \\ 0 & & \beta_m \end{Vmatrix}.$$

Donc,

$$A' \otimes B' = (C \otimes D)(A \otimes B)(C^{-1} \otimes D^{-1}) = (C \otimes D)(A \otimes B)(C \otimes D)^{-1}$$

est une matrice triangulaire à coefficients diagonaux $\alpha_i \beta_j$ qui sont des valeurs propres de la matrice $A' \otimes B'$ et, par conséquent, de la matrice $A \otimes B$. On a $\det(A \otimes B) = \prod_{i,j} \alpha_i \beta_j = (\prod_i \alpha_i)^m (\prod_j \beta_j)^n = (\det A)^m (\det B)^n$.)

4. En se servant de la formule (8) et des tables données aux § 2, n° 1 ; § 5, n° 2; § 5, n° 4, vérifier que la décomposition

$$\Phi^{(3)} \otimes \Phi^{(3)} \approx \Phi^{(1)} \dotplus \Phi^{(2)} \dotplus \Phi^{(3)}$$

est vraie pour le carré tensoriel de la représentation de dimension deux $\Phi^{(3)}$ du groupe symétrique S_3, et qu'il en est de même de la décomposition

$$\Phi^{(5)} \otimes \Phi^{(5)} \approx \Phi^{(1)} \dotplus \Phi^{(2)} \dotplus \Phi^{(3)} \dotplus \Phi^{(4)}$$

pour le carré tensoriel de la représentation de dimension deux $\Phi^{(5)}$ du groupe quaternionien Q_8.

5. *Représentations du produit direct de groupes.* Soient donnés deux groupes G et H, avec les représentations linéaires (Φ, V), (Ψ, W). Alors, en posant

$$(\Phi \otimes \Psi)(g \cdot h) = \Phi(g) \oplus \Psi(h),$$

où $g \cdot h$ est un élément du produit direct $G \times H$ des groupes G, H, nous ferons opérer $G \times H$ sur le produit tensoriel $V \otimes_{\mathbb{C}} W$; comme à l'ordinaire, on a

$$(\Phi(g) \otimes \Psi(h))(v \otimes w) = \Phi(g)v \otimes \Psi(h)w.$$

Vérifier que l'application ainsi définie

$$\Phi \otimes \Psi: G \times H \to \mathrm{GL}(V \otimes W)$$

est une représentation du groupe $G \times H$ de caractère $\chi_{\Phi \otimes \Psi} = \chi_\Phi \chi_\Psi$. Démontrer l'assertion suivante. Soient $\Phi^{(1)}, \ldots, \Phi^{(r)}$ (respectivement $\Psi^{(1)}, \ldots, \Psi^{(s)}$) toutes les représentations irréductibles du groupe G (respectivement, du groupe H). Alors, les représentations $\Phi^{(i)} \otimes \Psi^{(j)}$ du groupe $G \times H$ sont irréductibles, et les seules représentations irréductibles du groupe $G \times H$ sont $\Phi^{(i)} \otimes \Psi^{(j)}$, $1 \leqslant i \leqslant r$, $1 \leqslant j \leqslant s$.

6. Les formes xy, $x^n + y^n$ sont des invariants du groupe linéaire diédral de dimension deux

$$(D_n, \Phi) = \left\langle \begin{Vmatrix} \varepsilon & 0 \\ 0 & \varepsilon^{-1} \end{Vmatrix}, \quad \begin{Vmatrix} 0 & 1 \\ 1 & 0 \end{Vmatrix} \right\rangle, \quad \varepsilon^n = 1$$

(voir § 5, exercice 9). Montrer que tout autre invariant (entier) du groupe (D_n, Φ) a la forme d'un polynôme en xy, $x^n + y^n$.

7. Vérifier que le groupe quaternionien considéré dans sa représentation irréductible de dimension deux ne possède pas d'invariants quadratiques et cubiques. Que peut-on dire des formes x^2y^2, $x^4 + y^4$?

CHAPITRE 9

SUR LA THÉORIE DES CORPS, ANNEAUX ET MODULES

La reprise de l'étude des structures algébriques qui ont été déjà examinées précédemment, est motivée par des considérations suivantes. Premièrement, il semble souhaitable de compléter, dans une certaine mesure, nos renseignements sur les corps et sur les anneaux, en s'appuyant là, où cela est nécessaire, sur une base solide de la théorie des groupes. Deuxièmement, les résultats obtenus au cours du chapitre 8 sur les représentations des groupes se trouvent incorporés de façon naturelle dans la théorie générale des modules sur les anneaux, et il serait regrettable de ne pas en faire mention ne serait-ce que sous une forme succincte. La notion fondamentale de module est bien importante par elle-même et mérite une étude plus détaillée; à cet effet, nous recommandons au lecteur de consulter d'autres ouvrages.

§ 1. Extensions finies des corps commutatifs

1. Eléments primitifs et degrés des extensions.— Si F est un corps commutatif contenant P comme sous-corps, on dit aussi que F est une *extension* du corps P (voir chap. 4, § 4). Nous nous bornerons au début à examiner le cas le plus simple, où l'extension $F = P(\theta)$ est obtenue à partir du corps P par adjonction (à l'intérieur du corps donné F) d'un seul élément $\theta \in F$. On dit dans ce cas que $P(\theta)$ est une *extension simple* (ou *monogène*) du corps P, et θ est un *élément primitif* de cette extension. D'après son sens, $P(\theta)$ est un corps des quotients de l'anneau intègre $P[\theta]$. L'élément θ est *transcendant* sur P (voir chap. 5, § 2) si, et seulement si, l'extension $P(\theta)$ est isomorphe au corps des fractions rationnelles. Pourtant, si θ est un élément *algébrique*, alors $P(\theta) \cong P[X]/(f(X))$ (voir chap. 5, § 2, relation (9) et chap. 5, § 2, corollaire du théorème 5). Ici, $f(X)$ est un polynôme irréductible de degré $n > 0$, dont θ est une racine. Réciproquement, si $f \in P[X]$ est un polynôme irréductible, on construit de façon canonique (voir chap. 6, § 3) un corps commutatif F dans lequel f possède au moins une racine θ. De la construction il résulte que F est identifié avec l'ensemble des éléments

de la forme

$$a_0 + a_1\theta + \ldots + a_{n-1}\theta^{n-1}, \quad a_i \in P, \ n = \deg f.$$

Pour les éléments de l'anneau $P[\theta]$, cela est évident (effectuer la division euclidienne de $g(X)$ par $f(X)$ et introduire $X = \theta$) ; quant à la division dans $P[\theta]$, elle est effectuée comme suit : si $g(X) = a_0 + a_1X + \ldots + a_{n-1}X^{n-1}$, l'irréductibilité de f entraîne l'égalité P.G.C.D. $(f, g) = 1$ et l'existence des polynômes $u(X)$, $v(X)$ de degré $< n$, pour lesquels $fu + gv = 1$; d'où l'on déduit que $g(\theta)\,v(\theta) = 1$ et $1/g(\theta) = v(\theta)$. Le nombre n peut être considéré comme dimension de l'espace vectoriel

$$F = \langle 1, \theta, \ldots, \theta^{n-1} \rangle_P$$

sur P, avec les éléments de base $1, \theta, \ldots, \theta^{n-1}$.

Dans le cas d'une extension arbitraire $F \supset P$, il est aussi raisonnable de considérer F comme espace vectoriel sur P. Sa dimension $\dim_P F$ (peut-être infinie) sera désignée par $[F:P]$ et appelée *degré d'extension* de F sur P. Si $F = P(\theta)$, on dit aussi que $[F:P]$ est *degré d'élément primitif*. Il est clair que pour un élément transcendant $\theta \in F$ les éléments $1, \theta, \theta^2, \ldots$ sont linéairement indépendants sur P, et $[P(\theta): P] = \infty$. D'autre part, de ce qui précède on déduit l'assertion suivante :

THÉORÈME 1.— *Soit F une extension quelconque du corps commutatif P. Un élément $\theta \in F$ est algébrique sur P si, et seulement si, $[P(\theta): P] < \infty$. En outre, l'algébricité de θ entraîne l'égalité $P(\theta) = P[\theta]$.* ∎

Nous appellerons $K \supset F \supset P$ *tour d'extensions à deux étages*. Elle permet de parler de trois espaces vectoriels : K/P (K sur P), K/F (K sur F) et F/P (F sur P). Leurs dimensions sont liées par une relation analogue à celle qui existe entre les indices des sous-groupes.

THÉORÈME 2. — *Dans une tour d'extensions $K \supset F \supset P$, le degré $[K:P]$ est fini si, et seulement si, les degrés $[K:F]$ et $[F:P]$ sont finis. Dans le cas de leur finitude, on a la relation*

$$[K:P] = [K:F]\,[F:P].$$

DÉMONSTRATION.—En supposant finis $[K:F]$ et $[F:P]$, choisissons une P-base $f_1, \ldots, f_m$ de F/P et une F-base $e_1, \ldots, e_n$ de K/F. Alors tout élément $x \in K$ s'écrit sous la forme $x = \sum \alpha_j e_j$, avec $\alpha_j \in F$. A son tour $\alpha_j = \sum p_{ij} f_i$, avec $p_{ij} \in P$. Par conséquent, $x = \sum_{i,j} p_{ij} f_i e_j$, et nous voyons que mn éléments $f_i e_j$ engendrent linéairement K sur P. Supposons leur dépendance linéaire : $\sum_{i,j} p_{ij} f_i e_j = 0$ pour certains $p_{ij} \in P$. Alors

$$0 = \sum_{i,j} p_{ij} f_i e_j = \sum_j \Big(\sum_i p_{ij} f_i\Big) e_j \Rightarrow \sum_i p_{ij} f_i = 0 \Rightarrow p_{ij} = 0$$

pour tous les $i = 1, \ldots, m$; $j = 1, \ldots, n$, car $e_1, \ldots, e_n$ sont linéairement indépendants sur F, et $f_1, \ldots, f_m$ sont linéairement indépendants sur P. Par suite, mn éléments $f_i e_j$ forment une base de l'espace vectoriel K/P, et $[K:P] = nm = [K:F]\,[F:P]$.

Réciproquement, l'inégalité $[K:P] < \infty$ entraîne que $[F:P]$ est fini, car F/P est un sous-espace de l'espace K/P. Si $\{a_1, \ldots, a_r\}$ est une P-base de K, un élément arbitraire $x \in K$ sera une combinaison linéaire de $a_1, \ldots, a_r$ à coefficients dans P et, à plus forte raison, à coefficients dans F. Le nombre d'éléments parmi $a_1, \ldots, a_r$, qui sont linéairement indépendants sur F, ne peut que diminuer. Ainsi, $[K:F] < \infty$. ■

COROLLAIRE.— *Soit F une extension du corps commutatif P et soit A l'ensemble de tous les éléments de F qui sont algébriques sur P. Alors, A est un sous-corps de F contenant P.*

DÉMONSTRATION.— Tout élément $t \in P$ est racine du polynôme linéaire $X - t \in P[X]$, de sorte que $P \subset A$. Soient $u, v \in A$. D'après le théorème 1, on a $[P(u): P] < \infty$. L'élément v qui est algébrique sur P, le sera aussi sur $P(u)$, c'est-à-dire $[P(u, v): P(u)] = [P(u)(v): P(u)] < \infty$. D'après le théorème 2, on a $[P(u, v):P] = [P(u, v):P(u)]\,[P(u):P] < \infty$.

Puisque $u - v$, $uv \in P(u, v)$, de nouveau d'après le théorème 1, on a $u - v$, $uv \in A$, c'est-à-dire que A est un sous-anneau de F. Il est un corps commutatif, car

$$0 \neq u \in A \Rightarrow [P(u^{-1}):P] = [P(u):P] < 8. \ ■$$

On dit qu'une extension $F \supset P$ est *algébrique sur P* si tous les éléments de F sont algébriques sur P. Tout élément α de l'extension algébrique est racine d'un polynôme unitaire (c'est-à-dire dont le coefficient dominant est égal à 1) non nul $f \in P[X]$ dépendant de α. Si $f(\alpha) = 0$ et $g(\alpha) \neq 0$ pour tout $0 \neq g \in P[X]$ avec $\deg g < \deg f$, on appelle $f = f_\alpha$ *polynôme minimal de l'élément* α. Le polynôme minimal est irréductible sur P, il est univoquement défini et son degré coïncide avec le degré d'un élément α (souvent, le polynôme obtenu en multipliant le polynôme minimal par une constante est aussi appelé polynôme minimal). Toutes les racines distinctes du polynôme f_α sont considérées comme *conjuguées de* α. L'explication de cette terminologie est donnée plus loin dans le théorème 3. Si car $P = 0$, le nombre de racines distinctes coïncide avec $\deg f_\alpha$ (voir chap. 6, § 1), mais dans le cas général, il n'en est pas ainsi (voir exercices 4 et 5).

D'après les résultats obtenus, l'extension $F \supset P$ de degré fini $[F:P]$ est une *extension algébrique finie*, c'est-à-dire elle est obtenue à partir de P par l'adjonction d'un nombre fini d'éléments algébriques $\alpha_1, \ldots, \alpha_m$. Réciproquement, *toute extension algébrique finie $F = P(\alpha_1, \ldots, \alpha_m)$ est de degré fini.* En effet, $f_k(\alpha_k) = 0$, $1 \leqslant$

$\leqslant k \leqslant m$, $f_k \in P[X]$. L'élément α_k qui est algébrique sur P, le sera naturellement aussi sur $P(\alpha_1, \ldots, \alpha_{k-1})$. On a donc $[P(\alpha_1, \ldots, \alpha_k) : P(\alpha_1, \ldots, \alpha_{k-1})] < \infty$, et suivant le théorème 2,

$$[F:P] = [P(\alpha_1, \ldots, \alpha_m) : P] =$$

$$= \prod_{k=1}^{m} [P(\alpha_1, \ldots, \alpha_k) : P(\alpha_1, \ldots, \alpha_{k-1})] < \infty. \blacksquare$$

Dans de nombreux cas (et en particulier, lorsque car $P = 0$; voir exercice 13), l'extension algébrique finie est simple. Dans les cas que nous considérons ci-dessous, l'existence d'un élément primitif s'établit directement.

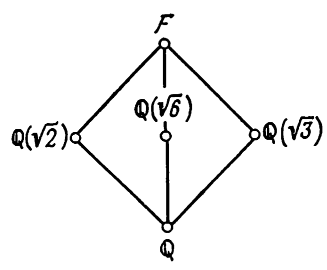
EXEMPLE. — Le corps commutatif $F = \mathbb{Q}(\sqrt{2}, \sqrt{3})$ considéré comme espace vectoriel sur $\mathbb{Q}$ est de dimension quatre: $F = \langle 1, \sqrt{2}, \sqrt{3}, \sqrt{6}\rangle_{\mathbb{Q}}$, c'est-à-dire que tout élément $\alpha \in F$ s'écrit sous la forme de la combinaison linéaire $\alpha = a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}$ à coefficients rationnels a, b, c, d. D'autre part, $F = \langle 1, \theta, \theta^2, \theta^3\rangle_{\mathbb{Q}}$, où $\theta = \sqrt{2} + \sqrt{3}$. En effet, $\sqrt{2} = -\frac{9}{2}\theta + \frac{1}{2}\theta^3$, $\sqrt{3} = \frac{11}{2}\theta - \frac{1}{2}\theta^3$, $\sqrt{6} = -\frac{5}{2} + \frac{1}{2}\theta^2$. L'élément primitif θ possède le polynôme minimal $f_\theta(X) = X^4 - 10X^2 + 1$ ayant pour racines

$\theta^{(1)} = \theta = \sqrt{2} + \sqrt{3}$, $\theta^{(2)} = \sqrt{2} - \sqrt{3}$, $\theta^{(3)} = -\sqrt{2} + \sqrt{3}$, $\theta^{(4)} = -\sqrt{2} - \sqrt{3}$.

On attire l'attention sur le fait que F est un corps de décomposition du polynôme $f_\theta(X)$ et que

$$F = \mathbb{Q}(\theta^{(1)}, \theta^{(2)}, \theta^{(3)}, \theta^{(4)}) = \mathbb{Q}(\theta^{(i)}),\ i = 1, 2, 3, 4.$$

Dans la théorie générale de Galois un tel corps serait appelé *corps normal*. Le diagramme des sous-corps du corps F



ressemble au diagramme des sous-groupes du groupe à quatre éléments V_4, et ce n'est pas un effet du hasard. Si nous considérons un automorphisme quelconque $\Phi: F \to F$ (voir chap. 4, § 4, n° 5), des relations $\Phi(x+y) = \Phi(x) + \Phi(y)$, $\Phi(xy) = \Phi(x)\Phi(y)$, $\forall x, y \in F$, il résulte que Φ est entièrement défini par son opération sur l'élément primitif θ. On a aussi $\Phi(a) = a$, $\forall a \in \mathbb{Q}$, et donc

$$\Phi(\theta)^4 - 10\Phi(\theta)^2 + 1 = \Phi(\theta^4 - 10\theta^2 + 1) = \Phi(0) = 0.$$

Par conséquent, $\Phi(\theta)$ est une des racines $\theta^{(i)}$, $i = 1, 2, 3, 4$, et nous arrivons à la conclusion que le groupe de tous les automorphismes Aut $(F/\mathbb{Q})$ que l'on appelle aussi *groupe de Galois* G $(F/\mathbb{Q})$ ou encore G (f_θ), a l'ordre $4 = [F:\mathbb{Q}]$. Les groupes d'ordre 4 ne sont, à un isomorphisme près, qu'au nombre de deux: le groupe cyclique Z_4 et $Z_2 \times Z_2 \cong V_4$. Des calculs directs montrent que Aut $(F/\mathbb{Q}) \cong V_4$.

On peut s'en rendre compte le plus facilement en considérant la représentation de Aut $(F/\mathbb{Q})$ par les permutations sur l'ensemble $\Omega = \{1, 2, 3, 4\}$ dont les éléments numérotent les racines $\theta^{(i)}$. Si par exemple, $\Phi(\theta^{(1)}) = \theta^{(2)}$, alors $\theta^{(1)}\theta^{(2)} = -1 \Rightarrow \theta^{(2)}\Phi(\theta^{(2)}) = -1 \Rightarrow \Phi(\theta^{(2)}) = \theta^{(1)}$ et $\Phi(\theta^{(3)}) = -\Phi(\theta^{(2)}) = -\theta^{(1)} = \theta^{(4)}$, c'est-à-dire $\Phi \approx (12)(34) = \sigma$. On obtient de la même manière les automorphismes $(13)(24) = \tau$ et $(14)(23) = \sigma\tau$.

Il reste à ajouter à ce qui vient d'être dit, que le sous-groupe cyclique $\langle\sigma\rangle$ laisse stables tous les éléments du sous-corps intermédiaire $\mathbb{Q}(\sqrt{2})$ et que $\langle\sigma\rangle$ est le groupe G $(F/\mathbb{Q}(\sqrt{2}))$ de tous les automorphismes (groupe de Galois) du corps commutatif F par rapport au sous-corps $\mathbb{Q}(\sqrt{2})$ De façon analogue $\mathbb{Q}(\sqrt{3})$ et $\mathbb{Q}(\sqrt{6})$ sont respectivement les corps des invariants pour $\langle\tau\rangle$ et $\langle\sigma\tau\rangle$ qui seront à leur tour les groupes de Galois G $(F/\mathbb{Q}\sqrt{(3)})$, G $(F/\mathbb{Q}(\sqrt{6}))$. Nous avons vérifié sur un exemple que la correspondance de Galois entre les sous-corps du corps normal F et les sous-groupes de son groupe des automorphismes est bijective.

2. Isomorphisme des corps de décomposition.— Au chapitre 6 § 3, où nous avons défini et construit un corps de décomposition F sur P du polynôme unitaire $f(X) = X^n + a_1X^{n-1} + \ldots$ $\ldots + a_n \in P[X]$, il a été indiqué que la construction de ce corps comportait des éléments d'arbitraire. En reprenant maintenant cette construction, nous pourrions dire seulement que $[F:P] \leqslant n!$ (essayez de comprendre pourquoi). Or, en réalité, tous les corps de décomposition sur P d'un polynôme donné f sont isomorphes. Afin de préciser cette proposition, analysons de plus près une situation plus générale.

Suivant le théorème 3 du chapitre 5, § 2, toute application isomorphe φ d'un corps P sur un corps $\widetilde{P}$ se prolonge de façon unique en un isomorphisme de $P[X]$ sur $\widetilde{P}[X]$, de sorte que

$$f(X) = X^n + a_1X^{n-1} + \ldots + a_n \mapsto \tilde{f}(X) = \varphi_X f = \\ = X^n + \varphi(a_1) X^{n-1} + \ldots + \varphi(a_n).$$

Théorème 3. — *Soit $\varphi: P \mapsto \widetilde{P}$ un isomorphisme des corps commutatifs, soient aussi $f \in P[X]$ un polynôme unitaire de degré $n > 0$, $\tilde{f} = \varphi_X f$ son image par l'isomorphisme φ_X; F, $\widetilde{F}$ les corps de décomposition des polynômes f, $\tilde{f}$ sur P et sur $\widetilde{P}$ respectivement. Alors φ se prolonge en un isomorphisme $\Phi: F \mapsto \widetilde{F}$ par $k \leqslant [F:P]$ procédés; ceci étant $k = [F:P]$ si toutes les racines du polynôme $\tilde{f}(X)$ sont distinctes.*

Démonstration. Etape I. — Considérons d'abord le cas des extensions quelconques $K \supset P$, $\widetilde{K} \supset \widetilde{P}$. Soit $\theta \in K$ un élément algébrique, avec un polynôme minimal $g = g_\theta \in P[X]$. On affirme que

l'isomorphisme $\varphi: P \to \tilde{P}$ se prolonge en monomorphisme $\rho: P(\theta) \to \tilde{K}$ si, et seulement si, $\tilde{g}$ possède une racine dans K et le nombre de prolongements coïncide avec le nombre de racines distinctes de $\tilde{g}$ dans $\tilde{K}$.

En effet, l'existence de ρ implique que l'élément $\rho(\theta)$ doit être racine de $\tilde{g}$: $g(\theta) = 0 \Rightarrow \tilde{g}(\rho(\theta)) = \rho(g(\theta)) = 0$. Réciproquement, si $\tilde{g}(\omega) = 0$, alors $\operatorname{Ker} \psi \supset g(X) P[X]$, où $\psi: P[X] \to \tilde{K}$ est un homomorphisme défini par la correspondance $u(X) \mapsto \tilde{u}(\omega)$. Comme dans le cas des groupes, ψ induit un homomorphisme $\bar{\psi}: P[X]/g(X) P[X] \to \tilde{K}$ (on a $u(X) + g(X) P[X] \mapsto \tilde{u}(\omega)$; si cela n'est pas tout à fait clair, il convient de se reporter aux résultats du § 2 qui suit). Remarquons que, $g(X)$ étant irréductible, l'anneau quotient $P[X]/g(X) P[X]$ est un corps de sorte que $\bar{\psi}$ est un monomorphisme. En procédant exactement de la même façon, on définit l'isomorphisme des corps $\bar{\sigma}: P[X]/g(X) P[X] \to P(\theta)$ $(u(X) + g(X) P[X] \mapsto u(\theta))$. La composée $\rho = \bar{\psi} \circ \bar{\sigma}^{-1}$ est une application monomorphe de $P(\theta)$ dans $\tilde{K}$ (vu que $\rho(u(\theta)) = \tilde{u}(\omega)$). Puisque $P(\theta)$ est engendré sur P par l'élément θ, il vient que ρ est l'unique prolongement de φ qui transforme θ en ω. Cela signifie justement que le nombre de monomorphismes différents ρ, avec la restriction $\rho|_P = \varphi$, est égal au nombre de racines distinctes de $\tilde{g}(X)$ dans $\tilde{K}$.

ETAPE II. — Le corps de décomposition a été construit par l'adjonction successive des racines des polynômes irréductibles. Utilisons ensuite la récurrence sur la dimension $[F:P]$.

Si $[F:P] = 1$, le polynôme f se décompose en facteurs linéaires déjà dans $P[X]$: $f(X) = (X - c_1) \ldots (X - c_n)$. Dans un tel cas, $\tilde{f}(X) = (\varphi_X f)(X) = (X - \tilde{c}_1) \ldots (X - \tilde{c}_n)$. Les racines $\tilde{c}_1, \ldots, \tilde{c}_n$ du polynôme $\tilde{f}$ sont contenues dans $\tilde{P}$, et puisque $\tilde{F}$ est engendré par ces racines sur $\tilde{P}$, on a $\tilde{F} = \tilde{P}$, de sorte que le prolongement $\Phi = \varphi_X$ est unique.

Si $[F:P] > 1$, décomposons $f(X)$ sur P en facteurs unitaires irréductibles, parmi lesquels il doit exister au moins un polynôme de degré $m > 1$. Désignons-le par $g(X)$. Puisque

$$f(X) = g(X) h(X) \Rightarrow \tilde{f}(X) = (\varphi_X f)(X) = \tilde{g}(X) \tilde{h}(X),$$

les polynômes $g(X)$ et $\tilde{g}(X)$ se décomposent sur les corps de décomposition F et $\tilde{F}$ de la façon suivante:

$$g(X) = (X - \theta_1) \ldots (X - \theta_m),$$
$$\tilde{g}(X) = (X - \omega_1) \ldots (X - \omega_m), \quad m \leqslant n.$$

Etant donnée son irréductibilité, $g(X)$ est le polynôme minimal de l'élément θ_1 sur P et $[P(\theta_1):P] = m$.

Si parmi $\omega_1, \ldots, \omega_m$ il y a l distincts, alors d'après l'étape I, il existe l applications monomorphes $\rho_1, \ldots, \rho_l$ de l'extension $L = P(\theta_1)$ dans $\widetilde{F}$, avec $\rho_i|_P = \varphi$. La structure du corps de décomposition est telle que F peut être considéré comme corps de décomposition sur L du polynôme $f \in L[X]$, et $\widetilde{F}$ comme corps de décomposition sur $\rho_i(L)$ du polynôme $\widetilde{f}(X)$, quel que soit $i = 1, 2, \ldots, l$. Suivant le théorème 2 on a l'inégalité $[F:L] = [F:P]/m < [F:P]$, si bien que par hypothèse de récurrence, chacun des ρ_i peut être prolongé en un isomorphisme $\Phi_{i,j}\colon F \to \widetilde{F}$, et le nombre de tels prolongements (le nombre d'indices j) n'est pas supérieur à $[F:L]$, il est égal à $[F:L]$ si toutes les racines dans $\widetilde{F}$ du polynôme $\widetilde{f}$ sont distinctes. Puisque $\Phi_{i,j}|_L = \rho_i$, $1 \leqslant j \leqslant [F:L]$, et $\rho_i|_P = \varphi$, alors $\Phi_{i,j}$ est un prolongement de φ, et $\rho_i \neq \rho_s \Rightarrow \Phi_{i,j} \neq \Phi_{s,t}$ pour $i \neq s$. Par suite, on obtient au total $k \leqslant m\,[F:L] = [F:P]$ prolongements de l'isomorphisme φ. Cette inégalité se transforme en une égalité si toutes les racines de $\widetilde{f}$ sont distinctes.

Etape III.— Soit, enfin, $\Phi\colon F \to \widetilde{F}$ un prolongement quelconque de φ. Tout comme dans l'étape II, la restriction $\Phi|_L$ qui est une application monomorphe de L dans $\widetilde{F}$, coïncide avec l'un des ρ_i, et dans ce cas Φ coïncide avec l'un des $\Phi_{i,j}$. ■

Corollaire 1.— *Deux corps de décomposition quelconques F, $\widetilde{F}$ sur P d'un polynôme $f \in P[X]$ sont isomorphes.*

En effet, il suffit de poser $\widetilde{P} = P$ dans le théorème 3 et de prendre pour φ l'application identique de P sur lui-même. ■

Corollaire 2. — *Le groupe des automorphismes* Aut (F/P) *de tout corps de décomposition F sur P d'un polynôme $f \in P[X]$ est un groupe fini d'ordre $\leqslant [F:P]$. Si toutes les racines du polynôme $f(X)$ sont distinctes,* $|\mathrm{Aut}(F/P)| = [F:P]$.

La démonstration découle immédiatement du théorème 3.

Remarque. — Bien que le corps de décomposition F sur $\mathbb{Q}$ (ou sur tout autre corps numérique) d'un polynôme $f \in \mathbb{Q}[X]$ puisse être considéré comme étant plongé dans le corps $\mathbb{C}$ des nombres complexes et donc par là même défini univoquement, le corollaire 2 montre que dans ce cas aussi on a eu intérêt à effectuer la démonstration pas trop agréable du théorème 3.

3. Corps commutatifs finis.— En plus de $Z_p = \mathbb{Z}/p\mathbb{Z}$, nous avons rencontré d'autres exemples de corps commutatifs finis (voir chap. 4, § 4). Il est temps de les inclure dans la théorie générale.

La première remarque évidente se rapporte à une extension finie arbitraire $K \supset F$ d'un corps commutatif fini F: *si* $|F| = q$ *et*

$[K:F] = n$, *alors* $|K| = q^n$. En effet, après le choix d'une base de l'espace vectoriel K/F, ce dernier s'identifie à l'espace F^n des vecteurs lignes $(\alpha_1, \ldots, \alpha_n)$ de longueur n. Toutes les coordonnées α_i prennent, indépendamment l'une de l'autre, q valeurs de F. Par conséquent, $|K| = |F^n| = q^n$. ■

Une deuxième remarque, liée à la première, consiste en ce que *tout corps commutatif fini F est de caractéristique finie p (p est un nombre premier), et $|F|$ est une puissance de p.* En effet, F étant fini, le sous-corps simple $P \subset F$ doit être isomorphe à un corps $Z_p = \mathbb{Z}/p\mathbb{Z}$. D'après la première remarque, l'extension finie $F \supset P$, avec $|P| = p$, est de puissance $|F| = p^m$. ■

Théorème 4. — *Pour tout corps commutatif fini F et pour tout entier positif n, il existe une et, à un isomorphisme près, une seule extension $K \supset F$ de degré $[K:F] = n$.*

Démonstration. a) Unicité. — Soit $K \supset F$ une extension de degré n. Comme nous le savons, $|F| = q \Rightarrow q = p^m$, où p est un nombre premier, et $|K| = q^n$. Par conséquent, le groupe multiplicatif $K^* = K \setminus \{0\}$ est d'ordre $q^n - 1$ et, suivant le théorème de Lagrange, l'ordre de chacun de ses éléments divise $q^n - 1 : t^{q^n-1} = 1$, $\forall t \neq 0$. Ceci signifie que tous les éléments du corps K (y compris $t = 0$) sont des racines distinctes du polynôme $X^{q^n} - X$ et que l'on a la décomposition

$$X^{q^n} - X = \prod_{t \in K} (X - t).$$

Une telle décomposition en produit de facteurs linéaires ne peut avoir lieu sur aucun sous-corps strict du corps K ayant un nombre d'éléments $< q^n$, et donc K est un corps de décomposition du polynôme $X^{q^n} - X$. En se reportant au corollaire 1 du théorème 3, on arrive à la conclusion requise.

b) Existence. — Les raisonnements développés dans la partie a) suggèrent une voie possible pour la construction de K. Prenons pour K le corps de décomposition sur F du polynôme $f(X) = X^{q^n} - X$. Puisque $q = p^m$, on a $q \cdot 1 = 0$ dans K. De ce fait, $f'(X) = q^n \cdot 1 \cdot X^{q^n-1} - 1 = -1$, et d'après le critère connu (chap. 6, § 1, théorème 4), $f(X)$ ne possède pas de racines multiples. Ceci signifie que le sous-ensemble $K_f \subset K$ des racines du polynôme $f(X)$ est de puissance $|K_f| = q^n$.

Puisque $K_f \subset K$ et car $K = p$, il vient suivant l'exercice 8 du chapitre 4, § 4 que $(x + y)^{p^s} = x^{p^s} + y^{p^s}$, quels que soient $x, y \in K_f$ $(F \subset K_f)$ et $s = 0, 1, 2, \ldots$ En particulier

$$x, y \in K_f \Rightarrow (x \pm y)^{q^n} = x^{q^n} \pm y^{q^n} = x \pm y \Rightarrow x \pm y \in K_f.$$

En outre,

$$1 \in K_f\,;\ \ (xy)^{q^n} = x^{q^n}y^{q^n} = xy \Rightarrow xy \in K_f\,;$$

$$0 \neq x \in K_f \Rightarrow (x^{-1})^{q^n} = x^{-1} \Rightarrow x^{-1} \in K_f.$$

Ainsi, K_f est un sous-corps de K, contenant F et toutes les racines du polynôme $f(X)$. Par la définition même du corps de décomposition, l'égalité $K_f = K$ doit être vérifiée. Le degré $[K:F]$ est égal à n car $q^{[K:F]} = |K| = |K_f| = q^n$. ■

Corollaire. — *Pour tout nombre premier p et pour tout entier positif n, il existe un et, à un isomorphisme près, un seul corps commutatif à nombre d'éléments p^n.*

La démonstration consiste à appliquer le théorème 4 au cas particulier où $|F| = p$. ■

Comme nous l'avons déjà dit au chapitre 4, § 4, on convient de désigner un corps commutatif fini à $q = p^n$ éléments par le symbole $\mathbb{F}_q$ ou encore, en l'honneur de E. Galois, par le symbole GF (p^n). Proposons-nous d'établir certaines propriétés des corps commutatifs finis.

Théorème 5. — *On a les assertions suivantes:*

(i) *Le groupe multiplicatif $\mathbb{F}_q^*$ d'un corps commutatif fini $\mathbb{F}_q$ est un groupe cyclique d'ordre $q - 1$.*

(ii) *Le groupe des automorphismes* Aut $(\mathbb{F}_q)$ *d'un corps commutatif fini $\mathbb{F}_q$ ayant un nombre d'éléments $q = p^n$, est un groupe cyclique d'ordre n et*

$$\mathrm{Aut}\,(\mathbb{F}_q) = \langle \Phi \mid \Phi(t) = t^p,\quad \forall t \in \mathbb{F}_q \rangle.$$

(iii) *Si $\mathbb{F}_{p^d}$ est un sous-corps d'un corps commutatif $\mathbb{F}_{p^n}$, alors $d \mid n$. Réciproquement, à tout diviseur d du nombre n correspond un sous-corps et un seul $\{t \in \mathbb{F}_{p^n} \mid \Phi^d(t) = t\} = \mathbb{F}_{p^d}$. Les automorphismes qui laissent fixes tous les éléments de ce sous-corps forment le groupe* Aut $(\mathbb{F}_{p^n}/\mathbb{F}_{p^d}) = \langle \Phi^d \rangle$. *Ainsi, il existe une bijection entre les sous-corps du corps commutatif fini $\mathbb{F}_q$ et les sous-groupes de son groupe des automorphismes (correspondance de Galois).*

(iv) *Si $q = p^n$ et $\mathbb{F}_q^* = \langle \theta \rangle$, alors θ est un élément primitif du corps commutatif, avec le polynôme minimal $h(X)$ de degré n. $\mathbb{F}_q$ est un corps de décomposition sur $\mathbb{F}_p$ du polynôme $h(X)$.*

(v) *Pour tout entier naturel m, il existe au moins un polynôme de degré m irréductible sur $\mathbb{F}_q$.*

Démonstration. (i) — Nous fournirons la démonstration d'une assertion plus générale. Soient F un corps commutatif quelconque et A un sous-groupe fini du groupe multiplicatif F^*. Au groupe abélien fini A sont applicables les résultats obtenus au chapitre 7, § 5. En particulier, nous savons que, si A est cyclique, ceci équivaut à dire que l'ordre $|A|$ coïncide avec l'exposant m du groupe A, qui est

le plus petit des entiers naturels pour lequel $a^m = 1$, $\forall a \in A$. Pour $m < |A|$, le polynôme $X^m - 1$ aurait dans F plus de m racines, ce qui est impossible. Par conséquent, A est un groupe cyclique.

(ii) — Considérons $\mathbb{F}_q$ comme extension finie $\mathbb{F}_q \supset \mathbb{F}_p$ de degré n de son sous-corps simple $\mathbb{F}_p \cong Z_p$. Puisque $\mathbb{F}_q$ est un corps de décomposition du polynôme $X^q - X$ dont toutes les racines sont distinctes, on a d'après le corollaire 2 du théorème 3, $|\text{Aut}(\mathbb{F}^q)| = n$. Les relations $(x + y)^p = x^p + y^p$, $(xy)^p = x^p y^p$, $1^p = 1$, indiquées au cours de la démonstration du théorème 4, entraînent que l'application $\Phi: t \mapsto t^p$ est un automorphisme du corps $\mathbb{F}_q$ (la condition de $\mathbb{F}_q$ fini est nécessaire). Si $\Phi^s: t \mapsto t^{p^s}$ est l'automorphisme unité, alors $t^{p^s} - t = 0$ pour tout $t \in \mathbb{F}_q$, d'où on déduit l'inégalité $s \geqslant n$. Or, pour $s = n$ nous obtenons réellement l'automorphisme unité, de sorte que $|\langle \Phi \rangle| = n$ et $\langle \Phi \rangle = \text{Aut}(\mathbb{F}_q)$.

(iii) — Suivant la première remarque sur les corps commutatifs finis (voir début du numéro), $p^n = (p^d)^r$, où r est le degré d'extension de $\mathbb{F}_{p^n} \supset \mathbb{F}_{p^d}$. Donc $n = dr$. Réciproquement, pour tout $d \mid n$, introduisons le sous-ensemble $F = \{t \in \mathbb{F}_{p^n} \mid t^{p^d} = t\}$. Puisque $n = dr \Rightarrow p^n - 1 = (p^d)^r - 1 = (p^d - 1)\, s$, on a

$$X^{p^n-1} - 1 = X^{(p^d-1)s} - 1 = (X^{p^d-1} - 1)\, g(X),$$

$$X^{p^n} - X = (X^{p^d} - X)\, g(X).$$

$\mathbb{F}_{p^n}$ étant un corps de décomposition du polynôme $X^{p^n} - X$, il y a exactement p^d éléments de $\mathbb{F}_{p^n}$ qui seront racines du polynôme $X^{p^d} - X$. Ce sont justement eux qui forment le sous-ensemble F qu'on peut maintenant identifier avec $\mathbb{F}_{p^d}$. Ce raisonnement, dual du théorème 4, établit aussi l'unicité du sous-corps commutatif à p^d éléments.

Remarquons que d'après la construction

$$\mathbb{F}_{p^d} = \{t \in \mathbb{F}_{p^n} \mid \Phi^d(t) = t\}$$

est l'ensemble de tous les éléments qui restent fixes lors de l'opération de $\langle \Phi^d \rangle$. Puisque $\text{Aut}(\mathbb{F}_{p^n}/\mathbb{F}_p) = \langle \Phi \rangle$ est un groupe cyclique, il est immédiat que tout automorphisme Φ^l n'appartenant pas à $\langle \Phi^d \rangle$ n'opère pas sur $\mathbb{F}_{p^d}$ de façon identique (il suffit de faire agir Φ^l sur l'élément générateur du groupe $\mathbb{F}_{p^d}^*$). Or, ceci signifie justement que le groupe des automorphismes relatifs $\text{Aut}(\mathbb{F}_{p^n}/\mathbb{F}_{p^d})$ coïncide avec $\langle \Phi^d \rangle$. La phrase finale de l'assertion (iii) a le même sens que dans l'exemple du n° 1.

(iv) — Il est tout à fait évident que $\mathbb{F}_q = \mathbb{F}_p(\theta)$, $q = p^n$. Soit $h(X) = X^n + a_1 X^{n-1} + \ldots + a_n$ le polynôme minimal de l'élément primitif θ. Puisque les éléments du sous-corps simple $\mathbb{F}_p$ restent fixes lors de tous les automorphismes et $a_i \in \mathbb{F}_p$, les racines de $h(X)$ sont θ, θ^p, θ^{p^2}, ..., $\theta^{p^{n-1}}$. Elles appartiennent toutes à notre

corps, et $\mathbb{F}_p(\theta, \ldots, \theta^{p^{n-1}}) = \mathbb{F}_p(\theta) = \mathbb{F}_{p^n}$ est un corps de décomposition sur $\mathbb{F}_p$ du polynôme $h(X)$.

(v) — En partant du théorème 4, construisons une extension $K \supset \mathbb{F}_q$ de degré m. D'après (i), K^* est un groupe cyclique. Si $K^* = \langle\theta\rangle$ et $h(X)$ est le polynôme minimal de l'élément primitif θ, alors $K = \mathbb{F}_q(\theta)$ et $\deg h(X) = [\mathbb{F}_q(\theta): \mathbb{F}_q] = [K: \mathbb{F}_q] = m$. Le polynôme minimal est, par définition, irréductible (sur $\mathbb{F}_q$) et nous avons donc ce qu'il fallait démontrer. ■

Après quelques préparatifs assez simples relevant de la théorie des nombres, nous obtenons une formule exacte pour le nombre de polynômes de degré m irréductibles sur $\mathbb{F}_q$.

4. Formule d'inversion de Möbius et ses applications.— La fonction μ définie en Théorie des nombres par

$$\mu(n) = \begin{cases} 1 & \text{si } n=1, \\ (-1)^k & \text{si } n=p_1 \ldots, p_k,\ p_i \text{ sont des nombres premiers différents,} \\ 0 & \text{si } n \text{ est divisible par un carré} > 1, \end{cases}$$

porte le nom de *fonction de Möbius*. Il est clair que μ est une *fonction multiplicative* en ce sens que μ n'est pas identiquement nulle et que $\mu(nm) = \mu(n)\mu(m)$ quels que soient n et m premiers entre eux. Il est également clair que, si $n = p_1^{m_1} \ldots p_r^{m_r}$, alors $\sum_{d|n} \mu(d) = \sum_{d|n_0} \mu(d)$, où $n_0 = p_1, \ldots p_r$ est le diviseur maximal de n, ne contenant pas de carrés. A son tour, le nombre de diviseurs $d = p_{i_1} \ldots p_{i_s}$ du nombre n_0, à s fixe, est égal à $\binom{r}{s}$. Ainsi, pour $n > 1$, on a

$$\sum_{d|n} \mu(d) = \sum_{d|n_0} \mu(d) = \sum_{s=0}^{r} \binom{r}{s} (-1)^s = (1-1)^r = 0$$

(la sommation au premier membre est effectuée sur tous les diviseurs $d \geqslant 1$ du nombre entier n). Finalement, on obtient la formule

$$\sum_{d|n} \mu(d) = \begin{cases} 1 \text{ si } n=1, \\ 0 \text{ si } n>1. \end{cases} \tag{1}$$

Il est aussi utile d'avoir sa modification suivante:

$$\sum_{d|n|m} \mu\left(\frac{m}{n}\right) = \begin{cases} 1 \text{ si } d=m; \\ 0 \text{ si } d|m,\ d<m \end{cases} \tag{2}$$

(la sommation est effectuée sur tous les n divisant m, qui sont divisibles par d). En posant $m = dt$, $n = dl$ et en faisant l parcourir les diviseurs du nombre t, nous passerons facilement de (2) à (1) et inversement.

La formule (1) (ou (2)) pourrait être prise pour une définition inductive de la fonction de Möbius. Son importance est contenue dans l'assertion suivante. *Soient f et g deux fonctions arbitraires de* $\mathbb{N}$ *dans* M ($M = \mathbb{Z}$, $\mathbb{R}$, $F[X]$, etc.), *liées par la relation*

$$f(n) = \sum_{d\mid n} g(d). \tag{3}$$

Alors

$$g(n) = \sum_{d\mid n} \mu\left(\frac{n}{d}\right) f(d). \tag{4}$$

En effet, compte tenu de (2), une sommation directe sur n divisant m des deux membres de l'égalité (3) multipliés par $\mu\left(\frac{m}{n}\right)$, donne

$$\sum_{n\mid m} \mu\left(\frac{m}{n}\right) f(n) = \sum_{n\mid m} \mu\left(\frac{m}{n}\right) \cdot \sum_{d\mid n} g(d) = \sum_{d\mid m} g(d) \cdot \sum_{d\mid n\mid m} \mu\left(\frac{m}{n}\right) = g(m).$$

Un simple changement des désignations conduit à la formule (4), connue sous le nom de *formule d'inversion de Möbius*. Le passage de (4) à (3) s'opère d'une manière analogue. ■

Il existe encore un analogue multiplicatif de la formule d'inversion de Möbius. *Si*

$$f(n) = \prod_{d\mid n} g(d),$$

alors

$$g(n) = \prod_{d\mid n} f(d)^{\mu\left(\frac{n}{d}\right)}. \tag{5}$$

Pour la démonstration, il convient d'effectuer les mêmes calculs formels :

$$\prod_{n\mid m} f(n)^{\mu\left(\frac{m}{n}\right)} = \prod_{n\mid m} \prod_{d\mid n} g(d)^{\mu\left(\frac{m}{n}\right)} = \prod_{d\mid m} \prod_{d\mid n\mid m} g(d)^{\mu\left(\frac{m}{n}\right)} =$$

$$= \prod_{d\mid m} g(d)^{\sum\limits_{d\mid n\mid m} \mu\left(\frac{m}{n}\right)} = g(m),$$

et de modifier légèrement les désignations.

Ci-dessous, nous donnons quatre exemples d'application de la formule d'inversion de Möbius.

Exemple 1 (*fonction* φ *d'Euler*). — Par définition, $\varphi(n)$ est le nombre des entiers premiers avec n de la série $0, 1, \ldots, n-1$,

ou ce qui revient au même, $\varphi(n) = |U(Z_n)|$ est l'ordre du groupe des éléments inversibles de l'anneau $Z_n = \mathbb{Z}/n\mathbb{Z}$. Au chapitre 8, § 1, exercice 5, nous avons obtenu la relation

$$n = \sum_{d|n} \varphi(d). \tag{6}$$

En utilisant la formule (4), on obtient

$$\varphi(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) d = \sum_{d|n} \mu(d) \frac{n}{d} = n \sum_{d|n} \frac{\mu(d)}{d}.$$

Si $n = p_1^{m_1} \dots p_r^{m_r}$, alors

$$\sum_{d|n} \frac{\mu(d)}{d} = 1 - \sum_i \frac{1}{p_i} + \sum_{i<j} \frac{1}{p_i p_j} - \dots + (-1)^r \frac{1}{p_1 p_2 \dots p_r} =$$
$$= \left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right) \dots \left(1 - \frac{1}{p_r}\right).$$

Ainsi,

$$\varphi(n) = n\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right) \dots \left(1 - \frac{1}{p_r}\right).$$

De cette formule, que nous avons donnée encore au chapitre 1, § 8, exercice 3, il résulte directement que la fonction φ est multiplicative.

EXEMPLE 2 (*polynômes cyclotomiques*). — Le corps de décomposition Γ_n sur $\mathbb{Q}$ du polynôme $X^n - 1$ s'appelle *corps cyclotomique*. Puisque toutes les racines n-ièmes de l'unité forment un groupe cyclique d'ordre n, le corps cyclotomique est de la forme $\Gamma_n = \mathbb{Q}(\zeta)$, où ζ est l'une des racines primitives ($\zeta \in \mathbb{C}$). Cherchons le degré $[\Gamma_n : \mathbb{Q}]$ et le polynôme minimal de l'élément ζ sur $\mathbb{Q}$.

Désignons par le symbole P_n l'ensemble de puissance $|P_n| = \varphi(n)$ des racines primitives n-ièmes de l'unité. Les sous-groupes d'un groupe cyclique d'ordre n sont en correspondance bijective avec les diviseurs d du nombre n (chap. 4, § 3, théorème 6), et chaque racine ζ^i appartient à un ensemble P_d. On en déduit qu'il existe une partition en classes disjointes:

$$\{1, \zeta, \zeta^2, \dots, \zeta^{n-1}\} = \bigcup_{d|n} P_d \tag{7}$$

(en passant aux puissances des ensembles, nous aurions retrouvé la relation (6)). On appelle *polynôme cyclotomique* correspondant à Γ_n, le polynôme

$$\Phi_n(X) = \prod_{\varepsilon \in P_n} (X - \varepsilon)$$

de degré $\varphi(n)$. Conformément à la partition (7), on a la décomposition

$$X^n - 1 = \prod_{i=0}^{n-1} (X - \zeta^i) = \prod_{d|n} \Big\{ \prod_{\varepsilon \in P_d} (X - \varepsilon) \Big\} = \prod_{d|n} \Phi_d(X). \tag{8}$$

En appliquant à (8) la formule multiplicative d'inversion de Möbius (5), on obtient une expression explicite pour Φ_n :

$$\Phi_n(X) = \prod_{d|n} (X^d - 1)^{\mu\left(\frac{n}{d}\right)}. \tag{9}$$

Pour de faibles valeurs de n, on a

$$\Phi_1(X) = X - 1,\ \Phi_2(X) = X + 1,\ \Phi_3(X) = X^2 + X + 1,$$
$$\Phi_4(X) = X^2 + 1,\ \Phi_6(X) = X^2 - X + 1,\ \Phi_8(X) = X^4 + 1,$$
$$\Phi_9(X) = X^6 + X^3 + 1,\ \Phi_{10}(X) = X^4 - X^3 + X^2 - X + 1,$$
$$\Phi_{12}(X) = X^4 - X^2 + 1.$$

Remarquons que

$$\Phi_n(X) \in \mathbb{Z}[X] \text{ et } \Phi_n(0) = 1, \quad n > 1. \tag{10}$$

Pour obtenir (10), on peut, sans passer par (9), raisonner par récurrence. Pour de faibles valeurs de n, cela est vérifié. Raisonnons plus loin comme suit. Soit

$$g(X) = \prod_{d|n,\ d \neq n} \Phi_d(X)$$

un polynôme unitaire à coefficients entiers. En appliquant l'algorithme de division euclidienne (chap. 5, § 2, théorème 5), nous obtenons des polynômes univoquement définis q, $r \in \mathbb{Z}[X]$, tels que $X^n - 1 = q(X)\,g(X) + r(X)$, $\deg r(X) < \deg g(X)$. Mais $X^n - 1 = \Phi_n(X)\,g(X)$ dans $\mathbb{Q}[X]$, et nous voyons que $\Phi_n(X) = q(X) \in \mathbb{Z}[X]$, le polynôme $\Phi_n(X)$ étant unitaire par suite de la même propriété de $g(X)$.

On a un théorème disant que $\Phi_n(X)$ *est un polynôme irréductible sur* $\mathbb{Q}$ *et donc* $\Gamma_n = \mathbb{Q}(\zeta)$ *est une extension de degré* $\varphi(n)$, *avec le polynôme minimal* $\Phi_n(X)$ *pour* ζ. Nous n'allons pas démontrer ce théorème et rappelons seulement qu'au chapitre 5, à la fin du § 3, nous avons établi l'irréductibilité de $\Phi_p(X) = (X^p - 1)/(X - 1) = X^{p-1} + X^{p-2} + \ldots + 1$, où p est un nombre premier arbitraire.

Il y a lieu de signaler que les corps cyclotomiques qui ont joué un grand rôle dans l'histoire du développement de la théorie des nombres algébriques, attirent jusqu'à présent l'attention des chercheurs.

Exemple 3 (*polynômes irréductibles sur* $\mathbb{F}_q$). — Soit $\Psi_d(q)$ le nombre total de polynômes unitaires de degré d irréductibles sur $\mathbb{F}_q$, $q = p^n$, et soit $f(X)$ un de ces polynômes. Son corps de décomposition sur $\mathbb{F}_q$ est isomorphe aussi bien à l'anneau quotient $\mathbb{F}_q[X]/f(X)\,\mathbb{F}_q[X]$ qu'au corps de décomposition du polynôme $X^{q^d} - X$ (corollaire du théorème 4). L'existence d'une racine θ

commune aux polynômes $X^{q^d} - X$ et $f(X)$ entraîne, par suite de l'irréductibilité de $f(X)$ que $X^{q^d} - X$ est divisible par $f(X)$. Puisque $X^{q^d} - X$ est un diviseur du polynôme $X^{q^m} - X$ pour tout $m = rd$, et vu que $X^{q^m} - X$ ne possède pas de racines multiples, nous arrivons à la conclusion que dans la décomposition de $X^{q^m} - X$ sur $\mathbb{F}_q$ figurent tous les polynômes unitaires irréductibles

$$f_{d,1}(X),\ f_{d,2}(X),\ \ldots,\ f_{d,\,\Psi_d(q)}(X)$$

de tout degré $d \mid m$, chaque polynôme n'intervenant qu'une seule fois:

$$X^{q^m} - X = \prod_{d\mid m} \Big\{ \prod_{k=1}^{\Psi_d(q)} f_{d,k}(X) \Big\}. \tag{11}$$

Le calcul des degrés des polynômes figurant aux deux membres de l'égalité (11) nous conduit à la relation

$$q^m = \sum_{d\mid m} d\Psi_d(q),$$

qui permet d'obtenir, par une application directe de la formule d'inversion de Möbius (4), l'expression pour $\Psi_m(q)$:

$$\Psi_m(q) = \frac{1}{m} \sum_{d\mid m} \mu\left(\frac{m}{d}\right) q^d. \tag{12}$$

Soit, par exemple, $q = 2$. Alors

$$\Psi_2(2) = \frac{1}{2}(2^2 - 2) = 1, \qquad \Psi_3(2) = \frac{1}{3}(2^3 - 2) = 2,$$

$$\Psi_4(2) = \frac{1}{4}(2^4 - 2^2) = 3, \qquad \Psi_5(2) = \frac{1}{5}(2^5 - 2) = 6,$$

$$\Psi_6(2) = \frac{1}{6}(2^6 - 2^3 - 2 + 2) = 9$$

(comparer avec chap. 6, § 1, exercice 10). La formule (12) montre que la probabilité de ce qu'un polynôme unitaire de degré m sur $\mathbb{F}_q$, choisi au hasard, soit irréductible est voisine de $1/m$. Pourtant, il n'existe pas de critères satisfaisants qui permettent d'établir l'irréductibilité d'un polynôme concret. Que peut-on dire, par exemple, de l'irréductibilité du trinôme $X^m + X^k + 1$ sur $\mathbb{F}_2$? De telles questions se posent constamment dans la théorie algébrique du codage (chap. 1, § 2, problème 3) et lors de la construction des suites pseudo-aléatoires.

EXEMPLE 4 (*constructions à la règle et au compas*). — Soit $P \subset$ CS (voir chap. 5, § 1, exemple 2) un corps des nombres construits, qui est une extension finie du corps $\mathbb{Q}$. Supposons d'abord que P soit un corps purement réel, c'est-à-dire que ses éléments soient des nombres réels. En particulier, l'élément primitif $\theta \in P$ (voir exercice 13)

est un nombre réel qu'on peut construire (comme longueur d'un segment) avec un nombre fini de pas au moyen d'une règle et d'un compas. Cela signifie que θ est un élément du corps $\mathbb{Q}(\theta_1, \theta_2, \ldots, \theta_r)$ et $[\mathbb{Q}(\theta_1, \ldots, \theta_k) : \mathbb{Q}(\theta_1, \ldots, \theta_{k-1})] \leqslant 2$. La dernière inégalité est compréhensible vu que θ_k est solution des équations à coefficients dans $\mathbb{Q}(\theta_1, \ldots, \theta_{k-1})$ de deux droites, d'une droite et d'une circonférence ou de deux circonférences. Les résultats du n° 1 sur les degrés des extensions algébriques montrent que $[\mathbb{Q}(\theta_1, \ldots, \theta_r) : \mathbb{Q}] = 2^m$, où $m \leqslant r$. Comme $\mathbb{Q}(\theta) \subset \mathbb{Q}(\theta_1, \ldots, \theta_r)$, on a d'après le théorème 2 que $[\mathbb{Q}(\theta) : \mathbb{Q}]$ est égal à une puissance de 2.

En revenant à un corps P quelconque (non nécessairement réel), écrivons-le aussi sous la forme $P = \mathbb{Q}(\theta)$. Maintenant, l'élément primitif $\theta = a + ib$ est un nombre complexe à composantes réelles construites a, b. Si $f(X)$ est le polynôme minimal (à coefficients rationnels) pour θ, alors $f(\theta) = 0$ et $f(\bar{\theta}) = 0$, où $\bar{\theta} = a - ib$. Il est clair que $\mathbb{Q}(\theta, \bar{\theta})$ est une extension algébrique finie du corps $\mathbb{Q}$. Ses éléments $a = (\theta + \bar{\theta})/2$ et $ib = (\theta - \bar{\theta})/2$ sont algébriques sur $\mathbb{Q}$; l'élément $b = ib/i$ sera, lui aussi, algébrique (voir corollaire du théorème 2), car $i^2 + 1 = 0$.

Ainsi, $\mathbb{Q}(a, b)$ est une extension algébrique réelle et finie du corps $\mathbb{Q}$ à éléments construits a, b. Du fait de ce qui précède, $[\mathbb{Q}(a, b) : \mathbb{Q}] = 2^m$. L'irréductibilité de $X^2 + 1$ sur $\mathbb{Q}(a, b) \subset \mathbb{R}$ signifie que $[\mathbb{Q}(a, b)(i) : \mathbb{Q}(a, b)] = 2$ et $[\mathbb{Q}(a, b)(i) : \mathbb{Q}] = 2^{m+1}$. Comme $P = \mathbb{Q}(\theta) \subset \mathbb{Q}(a, b, i)$, $[\mathbb{Q}(\theta) : \mathbb{Q}]$ divise 2^{m+1}. Nous avons démontré l'assertion importante suivante:

Si le corps P des nombres construits est une extension albébrique finie du corps $\mathbb{Q}$, alors $[P : \mathbb{Q}] = 2^n$, avec n un entier non négatif.

Ce résultat permet de répondre à certaines questions dont s'occupaient encore les mathématiciens antiques.

a) Peut-on construire (au moyen d'une règle et d'un compas) l'arête d'un cube dont le volume est égal à 2 (problème hindou de duplication du cube)? On suppose que le cube de volume unité est donné. Le polynôme $X^3 - 2$ dont la racine est la valeur cherchée de l'arête, est irréductible sur $\mathbb{Q}$, de sorte que $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3 \neq 2^n$. La réponse à la question posée est négative.

b) Est-ce tout angle qu'on peut diviser au moyen d'une règle et d'un compas en trois parties égales (problème de trisection de l'angle)? La réponse est négative même pour un angle concret de 60°, car la possibilité de construire $\varphi = 20°$ aurait signifié celle de $\cos \varphi$ et de $2\cos \varphi$, or, il n'en est pas ainsi. En effet, d'après le théorème de Moivre on a $1/2 = \cos 60° = \cos 3\varphi = 4\cos^3 \varphi - 3\cos \varphi$, de sorte que $\theta = 2\cos \varphi$ est racine du polynôme $f(X) = X^3 - 3X - 1 \in \mathbb{Z}[X]$. Comme ± 1 ne sont pas racines du polynôme

$f(X)$, ce dernier est irréductible sur $\mathbb{Q}$ (voir chap. 6, § 4, excercice 8) et donc $[\mathbb{Q}(\theta):\mathbb{Q}] = 3 \neq 2^n$.

c) Des considérations analogues montrent qu'au moyen d'une règle et d'un compas on ne peut pas construire tout polygone régulier à n côtés, où n est un entier naturel quelconque. Par exemple pour $n = 7$, il n'est pas difficile de se convaincre que le nombre $\theta = 2\cos\frac{360°}{7}$ à construire est racine du polynôme $X^3 + X^2 - 2X - 1$ qui est irréductible sur $\mathbb{Q}$.

Dès le début de son activité mathématique, le Grand Gauss a trouvé les conditions nécessaires et suffisantes auxquelles doit satisfaire le nombre n pour que la construction d'un polygone régulier à n côtés soit réalisable. En particulier, il a établi que le nombre premier n doit être un nombre de Fermat : $n = 2^{2^k} + 1$. La résolution complète de ces questions est liée à l'étude du groupe de Galois d'un corps cyclotomique (voir exemple 2).

EXERCICES

1. Montrer qu'une extension $F \supset P$ de degré un nombre premier ne possède pas de sous-corps stricts ($\neq P$, F).

2. Trouver l'élément primitif de l'extension $\mathbb{Q}(\sqrt{p}, \sqrt{q})$, où p et q sont des nombres premiers.

3. Trouver la dimension sur $\mathbb{Q}$ du corps de décomposition du polynôme $X^p - 2$.

4. Montrer que pour le polynôme $X^p - a$ sur le corps commutatif P de caractéristique $p > 0$, il n'existe que deux possibilités: être irréductible ou bien être la puissance p-ième d'un polynôme linéaire. (Indication. Considérer le corps de décomposition F du polynôme $X^p - a$. Soit $\theta \in F$ l'une des racines, de sorte que $a = \theta^p$ et $X^p - a = (X - \theta)^p$. Si maintenant $X^p - a = u(X)v(X)$, où $u(X)$ est un polynôme unitaire sur P de degré positif $m < p$, alors, $F[X]$ étant factoriel, on doit avoir l'égalité $u(X) = (X - \theta)^m$. En particulier, θ^m, $\theta^p \in P \Rightarrow \theta \in P$.)

5. Soit $Z_p(Y)$ un corps commutatif des fractions rationnelles de caractéristique p. Montrer que le polynôme $X^p - Y$ est irréductible sur $Z_p(Y)$ et que toutes ses racines coïncident. (Indication. D'après l'exercice précédent, il suffit de s'assurer que l'égalité $X^p - Y = \left(X - \frac{g(Y)}{h(Y)}\right)^p$, avec $g, h \in Z_p[Y]$, est impossible.)

6. Démontrer que pour tout $d \mid n$, $d < n$, on a la relation $X^n - 1 = (X^d - 1)\Phi_n(X)h_d(X)$, où $h_d \in \mathbb{Z}[X]$. (Indication. D'après (8) on a $X^d - 1 = \prod_{e \mid d} \Phi_e(X)$. Donc,

$$X^n - 1 = (X^d - 1) \prod_{s \mid n;\ s \nmid d} \Phi_s(X) = (X^d - 1)\Phi_n(X) \prod_{s \mid n;\ s \nmid d;\ s \neq n} \Phi_s(X).$$

Il ne reste qu'à se référer à (10).)

7. Soit q un entier positif > 1. D'après (10), $\Phi_n(q) \in \mathbb{Z}$. Montrer que $\Phi_n(q) \mid (q - 1) \Rightarrow n = 1$. (Indication. Comme $\Phi_n(X) = \prod (X - \varepsilon)$,

où ε parcourt les racines primitives, il vient que pour $n > 1$ on a tous les $\varepsilon \neq 1$, et de ce fait, la distance sur le plan $\mathbb{C}$ du point q à tout ε est plus grande que celle de q à 1. Par conséquent, $|\Phi_n(q)| = \prod (q - \varepsilon) > q - 1$, et $q - 1$ ne peut aucunement être divisible par $\Phi_n(q)$.)

8. Vérifier que le polynôme cyclotomique $\Phi_{15}(X) = X^8 - X^7 + X^5 - X^4 + X^3 - X + 1$, considéré sur le corps $\mathbb{F}_2$, est le produit de deux polynômes irréductibles $X^4 + X^3 + 1$ et $X^4 + X + 1$. En se servant de ce fait, démontrer que $\Phi_{15}(X)$ est irréductible sur $\mathbb{Q}$ (comparer avec chap. 6, § 1, exercice 11).

9. En partant de la suite d'inclusions naturelles

$$\mathrm{GF}(p) \subset \mathrm{GF}(p^{2!}) \subset \mathrm{GF}(p^{3!}) \subset \ldots,$$

introduire un *corps* dit *limite* $\Omega_p = \mathrm{GF}(p^{\infty!})$ en posant $\alpha \in \Omega_p \Leftrightarrow \{\alpha \in \mathrm{GF}(p^{n!})$, avec n suffisamment grand$\}$. En s'appuyant sur les propriétés fondamentales des corps finis, démontrer que Ω_p est un corps algébriquement clos. C'est ainsi qu'on obtient, compte tenu du corps $\mathbb{C}$ des nombres complexes, des exemples de corps algébriquement clos de toute caractéristique voulue.

10. Soit $q = p^n$. Montrer que pour $p = 2$, tous les éléments du corps commutatif $\mathbb{F}_q$ sont des carrés, et que pour $p > 2$, les carrés des éléments du groupe $\mathbb{F}_q^*$ forment un sous-groupe $\mathbb{F}_q^{*2}$ d'indice 2, et $\mathbb{F}_q^{*2} = \mathrm{Ker}\,(t \mapsto t^{(q-1)/2})$.

11. (M. Aschbacher). *Soit $\mathbb{F}_q$ un corps commutatif fini à nombre impair d'éléments $q = p^n$. Si $q \neq 3$ ou 5, il existe sur la « circonférence » $x^2 + y^2 = 1$ un point de coordonnées $x, y \in \mathbb{F}_q^*$.* Démontrer cette assertion pour $p > 5$. (Indication. Passer à l'équation $x^2 + y^2 - z^2 = 0$, avec $x, y, z \in \mathbb{F}_q$. D'après le théorème de Chevalley (voir chap. 6, § 1, exercice 4) le nombre total N de solutions de cette équation est divisible par p. Supposer que les solutions, avec $xyz \neq 0$, n'existent pas. Alors, calculer N en considérant deux cas. S'il n'existe pas de $a \in \mathbb{F}_p$, avec $a^2 + 1 = 0$, alors les seules solutions seront $(0, 0, 0)$, $(0, n, \pm n)$, $(n, 0, \pm n)$, $n = 1, 2, \ldots, p - 1$, et donc $N = 4p - 3 \equiv 0 \pmod{p} \Rightarrow p = 3$. Si $a^2 + 1 = 0$ pour un $a \in \mathbb{F}_p$, alors $N = 6p - 5 \equiv 0 \pmod{p} \Rightarrow p = 5$.)

12. Tout élément primitif d'un corps commutatif $\mathbb{F}_q$ est-il générateur du groupe multiplicatif $\mathbb{F}_q^*$? (Réponse: en général, non.)

13. (*Théorème sur l'élément primitif*). Soit $F = P(\theta_1, \theta_2, \ldots, \theta_r)$ une extension algébrique finie d'un corps commutatif P de caractéristique nulle. Montrer que $F = P(\theta)$ pour un élément θ algébrique sur P. (Indication. Par récurrence sur r, le problème se ramène au cas de $F = P(\alpha, \beta)$, où les éléments α et β sont algébriques sur P et ont les polynômes minimaux différents $f(X)$ et $g(X)$. Soit K un corps de décomposition du polynôme $f(X)\,g(X)$, si bien que l'on a

$$f(X) = (X - \alpha_1)(X - \alpha_2)\ldots(X - \alpha_n), \qquad \alpha_i \in K\ (\alpha_1 = \alpha),$$
$$g(X) = (X - \beta_1)(X - \beta_2)\ldots(X - \beta_m), \qquad \beta_j \in K\ (\beta_1 = \beta).$$

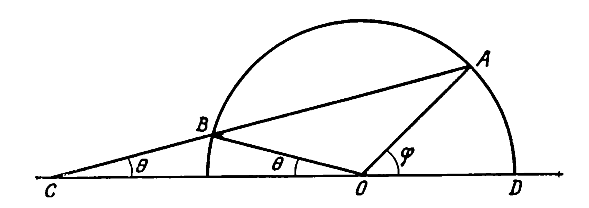
L'irréductibilité de $f(X)$ et de $g(X)$ ainsi que la condition car $P = 0$ nous garantissent (voir chap. 6, § 1, nº 4) que les éléments α_i, β_j sont deux à deux distincts, et nous sommes donc en mesure de construire des éléments $(\beta_j - \beta)/(\alpha - \alpha_i) \in K$, $i \neq 1$. Prenons un nombre rationnel quelconque $c \neq 0$ différent de ces éléments (de nouveau la condition car $P = 0$ est nécessaire!) et posons

$$\theta = \beta + c\alpha.$$

Il est clair que $P(\theta) \subset P(\alpha, \beta) = F$. Les polynômes $f(X)$ et $h(X) = g(\theta - cX) \in P(\theta)[X]$ possèdent une racine commune α. Si α_i est aussi leur racine commune dans K pour un $i > 1$, alors $0 = h(\alpha_i) = g(\theta - c\alpha_i)$, d'où

$\theta - c\alpha_i = \beta_j$ pour un $j \geqslant 1$. Or, cela signifie que : soit $c(\alpha - \alpha_i) = 0$, soit $c = (\beta_j - \beta)/(\alpha - \alpha_i)$. Les deux possibilités sont exclues du fait du choix de c, et par suite, $X - \alpha$ est le plus grand commun diviseur des polynômes $f, h \in K[X]$. Or, en réalité $f, h \in P(\theta)[X]$ et donc (voir chap. 5, § 3, nº 3) P.G.C.D $(f, h) \in P(\theta)[X]$. Par conséquent, $X - \alpha \in P(\theta)[X]$, c'est-à-dire $\alpha \in P(\theta)$ et $\beta = \theta - c\alpha \in P(\theta)$. Dans ce cas, $P(\alpha, \beta) \subset P(\theta)$ et donc $P(\alpha, \beta) = P(\theta)$.

14. La figure



montre l'un des procédés de trisection de l'angle : $\theta = \varphi/3$. La longueur des segments OB et CB est égale à 1. Le point A étant donné, comment peut-on construire le point C?

15. En construisant un corps commutatif concret $P \subset$ CS, montrer que le degré [CS : $\mathbb{Q}$] est infini.

§ 2. Quelques résultats relatifs aux anneaux

Ce paragraphe peut être considéré comme un complément fort utile, bien que petit, des chapitres 4 et 5.

1. **Nouveaux exemples d'anneaux factoriels.**— Au chapitre 5, § 3, nous avons vu que les anneaux euclidiens sont factoriels. Les anneaux $\mathbb{Z}$ et $P[X]$ en fournissent des exemples connus. Dans ce qui suit, nous donnons encore un exemple d'anneau euclidien ainsi qu'un exemple d'anneau factoriel qui n'est pas euclidien.

Exemple 1 (*anneau des entiers de Gauss*).— On a en vue **l'anneau** numérique

$$\mathbb{Z}[i] = \{m + in \mid m, \ n \in \mathbb{Z}\},$$

contenu dans le corps quadratique $\mathbb{Q}(i) \subset \mathbb{C}$, $i^2 + 1 = 0$, et identifié géométriquement à l'ensemble des nœuds (des points) d'un treillis à valeurs entières sur le plan $\mathbb{C}$. Il est clair que $\mathbb{Z}[i]$ est un anneau intègre. On définit sur $\mathbb{Z}[i]^* = \mathbb{Z}[i] \setminus \{0\}$ une application $\delta : \mathbb{Z}[i]^* \to \mathbb{N} \cup \{0\}$ en posant $\delta(m + in) = |m + in|^2 = m^2 + n^2$ (en d'autres termes, $\delta(a) = N(a)$ est la norme du nombre a dans $\mathbb{Q}(i)$; voir chap. 5, § 1, nº 5). On sait que $\delta(ab) = \delta(a)\,\delta(b) \geqslant \delta(a)$, quels que soient $a, b \in \mathbb{Z}[i]^*$, si bien que la propriété (E1) intervenant dans la définition de l'anneau euclidien (voir chap. 5, § 3, nº 3) est automatiquement satisfaite. Pour vérifier la validité de (E2), écrivons la fraction ab^{-1}, avec $b \neq 0$, sous la forme $ab^{-1} = \alpha + i\beta$, avec $\alpha, \beta \in \mathbb{Q}$, et choisissons les entiers k, l les plus pro-

ches de α, β, de manière que $\alpha = k + \nu$, $\beta = l + \mu$, $|\nu| \leqslant \frac{1}{2}$, $|\mu| \leqslant \frac{1}{2}$. Alors,

$$a = b\,[(k + \nu) + i\,(l + \mu)] = bq + r,$$

où $q = k + il \in \mathbb{Z}[i]$ et $r = b\,(\nu + i\mu)$. Puisque $r = a - bq$, on a $r \in \mathbb{Z}[i]$ et

$$\delta(r) = |r|^2 = |b|^2(\nu^2 + \mu^2) \leqslant \delta(b)\left(\frac{1}{4} + \frac{1}{4}\right) = \frac{1}{2}\,\delta(b) < \delta(b).$$

Par suite, $\mathbb{Z}\,[i]$ est un *anneau euclidien.* ■

L'anneau des entiers de Gauss $\mathbb{Z}\,[i]$ s'avère assez commode pour illustrer en miniature les méthodes utilisées en Théorie des nombres algébriques. C'est pourquoi nous allons examiner un peu plus en détail les propriétés de $\mathbb{Z}\,[i]$. Commençons par faire quelques remarques générales.

1) Un anneau intègre K dont tous les idéaux sont principaux, c'est-à-dire de la forme xK, s'appelle *anneau principal. Tous les anneaux euclidiens sont des anneaux principaux.* Pour $\mathbb{Z}$ et $P\,[X]$ ceci a été établi plus haut (voir chap. 5, § 2, corollaire du théorème 5) et, dans le cas général, la démonstration est tout à fait analogue: si J est un idéal d'un anneau euclidien K, alors $J = aK$ dès que $a \in J$ et $\delta\,(a) \leqslant \delta\,(x)$ pour tout $0 \neq x \in J$. ■

2) *Soient K un anneau euclidien arbitraire, avec fonction δ* (voir chap. 5, § 3, n° 3), *et $U\,(K)$ le groupe de ses éléments inversibles. Alors on a*

$$u \in U\,(K) \Leftrightarrow \delta\,(u) = \delta\,(1) \Leftrightarrow \delta\,(ux) = \delta\,(x), \quad \forall\, x \in K^*. \quad (1)$$

En effet, en vertu de (E1), $\delta\,(x) = \delta\,(1 \cdot x) \geqslant \delta\,(1)$ pour tout $x \in K^*$, et si $u \in U\,(K)$, alors $\delta\,(1) = \delta\,(u \cdot u^{-1}) \geqslant \delta\,(u)$, de sorte que $\delta\,(u) = \delta\,(1)$; on démontre de façon analogue que $u \in U\,(K) \Rightarrow \delta\,(ux) = \delta\,(x)$, $\forall x \in K^*$. Réciproquement, conformément à la remarque 1) on a $\delta\,(ux) = \delta\,(x) \Rightarrow uxK = xK$, $\forall x \in K^*$, ce qui implique $x = uxv \Rightarrow uv = 1 \Rightarrow u \in U\,(K)$. ■

Etant appliqué à l'anneau $\mathbb{Z}[i]$, le critère (1) signifie que $m + in \in U\,(\mathbb{Z}[i]) \Leftrightarrow m^2 + n^2 = 1$. *Par conséquent,* $U\,(\mathbb{Z}[i]) = \langle i \rangle$ *est un groupe cyclique d'ordre* 4.

3) Un idéal J de l'anneau K est dit *maximal*, si $J \neq K$ et tout idéal T contenant J coïncide avec K ou J. *Dans un anneau euclidien K, dire qu'un élément $p \in K$ est premier revient à dire que l'idéal pK est maximal.*

En effet, soient p un élément premier et $pK \subset T \subset K$, où T est un idéal de K. D'après 1) il vient que $T = aK$, et comme $p \in T$, on a $p = ab$, où l'un des éléments a, b est inversible. Si $a \in U\,(K)$, alors $T = aK = K$. Si $b \in U\,(K)$, alors $T = aK = abK = pK$. Réciproquement, supposons que l'idéal pK soit maximal et $p = ab$,

avec $a \notin U(K)$. Alors $aK \neq K$ et $pK \subset aK \Rightarrow pK = aK \Rightarrow a = pu = abu \Rightarrow bu = 1 \Rightarrow b \in U(K) \Rightarrow p$ est un élément premier. ∎

Examinons maintenant ce qu'un nombre premier $p \in \mathbb{Z}$ devient dans l'anneau $\mathbb{Z}[i]$. Il n'est pas exclu que p est encore élément premier dans $\mathbb{Z}[i]$. S'il n'en est pas ainsi, soit $p = \prod_{k=1}^{r} p_k$ son unique décomposition (d'après chap. 5, § 3, théorème 4) en facteurs premiers p_k dont le nombre est $r > 1$. En vertu de 2), $\delta(p_k) > 1$, si bien que $p^2 = \delta(p) = \Pi\delta(p_k)$ et le fait que $\mathbb{Z}$ est factoriel entraînant nécessairement les égalités $r = 2$, $p = p_1p_2$, $\delta(p_1) = \delta(p_2) = p$. Si $p_1 = m + in$, alors $p = \delta(p_1) = m^2 + n^2 = (m + in)(m - in) \Rightarrow p_2 = m - in$. Ainsi, *lorsque un nombre premier $p \in \mathbb{Z}$ admet une décomposition non triviale dans $\mathbb{Z}[i]$, on a*

$$p = (m + in)(m - in) = m^2 + n^2, \tag{2}$$

où $m + in$, $m - in$ sont des éléments premiers de $\mathbb{Z}[i]$. ∎

En particulier, $2 = (1 + i)(1 - i)$ n'est pas un élément premier de $\mathbb{Z}[i]$. Remarquons ensuite que $t^2 \equiv 0$ ou $1 \pmod{4}$ pour tout $t \in \mathbb{Z}$. Aussi, pour un nombre premier impair p, qui n'est pas premier dans $\mathbb{Z}[i]$, le critère (2) conduit-il à la conclusion que

$$p = m^2 + n^2 \equiv 0{,}1 \text{ ou } 2 \pmod{4} \Rightarrow p = 4k + 1.$$

Dans le cas où $p = 4k + 1$, posons $t = (2k)!$. Comme on a évidemment $t = (-1)^{2k}(2k)! = (-1)(-2)\ldots(-2k) \equiv (p - 1) \times (p - 2)\ldots(p - 2k) \equiv ((p + 1)/2)\ldots(p - 2)(p - 1) \pmod{p}$, alors

$$t^2 \equiv (2k)!\,((p + 1)/2)\ldots(p - 2)(p - 1) \equiv (p - 1)! \pmod{p},$$

ou, compte tenu du théorème de Wilson (voir chap. 6, fin du § 1), $t^2 + 1 \equiv 0 \pmod{p}$. Si maintenant p est un élément premier de $\mathbb{Z}[i]$, il résulte de l'égalité $(t + i)(t - i) = t^2 + 1 = lp$, $l \in \mathbb{Z}$, en vertu du théorème 1 du chapitre 5, § 3, que p divise l'un des éléments $t + i$, $t - i$. Or, $t \pm i = p(m + in) \Rightarrow \pm 1 = pn$, $n \in \mathbb{Z}$, ce qui est manifestement impossible. Nous avons ainsi démontré l'assertion suivante:

Un nombre premier $p \in \mathbb{Z}$ reste premier dans $\mathbb{Z}[i]$ si, et seulement si, $p = 4k - 1$.

Tout nombre premier $p = 4k + 1$ est représentable sous la forme $p = m^2 + n^2$, où $m, n \in \mathbb{Z}$. ∎

On en déduit assez facilement le théorème général relevant de la théorie des nombres:

THÉORÈME 1. — *Un nombre $t \in \mathbb{Z}$ est représentable sous la forme d'une somme des carrés de deux nombres m, $n \in \mathbb{Z}$ si, et seulement si,*

chaque diviseur premier $p = 4k - 1$ de t apparaît avec un exposant pair dans la décomposition canonique de t en facteurs premiers.

En effet, compte tenu des faits déjà connus, il suffit de prouver que, *si* P.G.C.D. $(m, n) = 1$ *et* $p \mid (m^2 + n^2)$, *alors* $p = 4k + 1$. Ceci est suffisamment clair: P.G.C.D. $(m, n) = 1$, $m^2 + n^2 \equiv 0$ (mod. p) $\Rightarrow mn \not\equiv 0$ (mod. p) $\Rightarrow m^{p-1} \equiv 1$ (mod. p), $n^2 \equiv -m^2$ (mod. p) $\Rightarrow (m^{p-2}n)^2 = m^{2p-4}n^2 \equiv -m^{2p-2} \equiv -1$ (mod. p). Ainsi, il existe un entier $s \in \mathbb{Z}$, tel que $s^2 \equiv -1$ (mod. p), $s^4 \equiv 1$ (mod. p). Par conséquent, l'ordre $p - 1$ du groupe multiplicatif $\mathbb{Z}_p^*$ est divisible par 4, et $p = 4k + 1$. ■

D'après la remarque 3), dire que $p = 4k - 1$ est premier dans $\mathbb{Z}[i]$ revient à dire que l'idéal $p\,\mathbb{Z}\,[i]$ est maximal, ce qui s'exprime à son tour par la propriété de l'anneau quotient $\mathbb{Z}\,[i]/p\mathbb{Z}[i]$ d'être un corps à p^2 éléments (voir à ce propos chap. 4, § 4, exercice 14 et théorèmes d'isomorphie des anneaux au n° 2). Ce n'est pas étonnant si l'on tient compte du fait que pour $p = 4k - 1$ le polynôme $X^2 + 1$ considéré sur Z_p est irréductible.

EXEMPLE 2 (*extensions polynomiales des anneaux factoriels*). — Nous allons montrer que les anneaux des polynômes $\mathbb{Z}[X_1, \ldots, X_n]$ et $P\,[X_1, \ldots, X_n]$ (P est un corps commutatif) sont factoriels quel que soit n. Cette assertion importante résulte immédiatement du théorème suivant:

THÉORÈME 2. — *Si K est un anneau factoriel, il en est de même de l'anneau des polynômes $K\,[X]$.*

DÉMONSTRATION.— Elle est basée sur les propriétés des anneaux des polynômes, qui sont associées au lemme de Gauss (voir chap. 5, § 3). A savoir nous aurons besoin de deux propriétés suivantes:

a) *Les polynômes primitifs f, $g \in K\,[X]$ associés dans $Q(K)[X]$* ($Q\,(K)$ est le corps des quotients de l'anneau factoriel K) *sont associés dans $K\,[X]$* (un exercice facile).

b) *Un polynôme $f \in K\,[X]$ de degré positif, irréductible sur K est aussi irréductible sur $Q\,(K)$* (la démonstration donnée au chap. 5, § 3 pour $K = \mathbb{Z}$ est aussi valable dans le cas général).

En passant à la démonstration du théorème, écrivons le polynôme $f \in K\,[X]$ de degré positif sous la forme $f = \mathrm{d}\,(f)\,f_0$, où $\mathrm{d}\,(f)$ est le contenu de f, et f_0 est un polynôme primitif. Par récurrence sur le degré des polynômes primitifs, on obtient une décomposition de f_0 en un produit $f_0 = f_1 \ldots f_s$ des polynômes primitifs $f_1, \ldots, f_s$ irréductibles sur K. Soit $f_0 = g_1 \ldots g_t$ une autre décomposition de ce type. Alors, d'après b), f_i et g_j sont irréductibles sur $Q\,(K)$. Comme l'anneau $Q\,(K)\,[X]$ est factoriel (chap. 5, § 3, corollaire du théorème 4), on a $s = t$ et, à condition que les facteurs sont convenablement ordonnés, le polynôme f_i est associé à g_i dans $Q\,(K)\,[X]$ et donc (d'après a)) aussi dans $K\,[X]$. Quant au polynôme

f dont d (f) n'admet pas d'inverse dans K, on obtient sa décomposition en prenant la décomposition d $(f) = p_1 \ldots p_r$ en facteurs premiers $p_i \in K$. L'unicité d'une telle décomposition de f (en sens usuel) résulte de l'unicité de la décomposition de f_0 qui vient d'être établie, et du fait que K est un anneau factoriel, ce qui assure l'unicité de la décomposition d $(f) = p_1 \ldots p_r$. ■

On a des inclusions strictes:

$$\left\{ \begin{array}{c} \textit{anneaux} \\ \textit{euclidiens} \end{array} \right\} \subset \left\{ \begin{array}{c} \textit{anneaux} \\ \textit{principaux} \end{array} \right\} \subset \left\{ \begin{array}{c} \textit{anneaux} \\ \textit{factoriels} \end{array} \right\}. \quad (3)$$

Quant à la première inclusion, nous la connaissons (voir remarque (1)). Il existe des exemples (nous ne les donnons pas ici) qui montrent qu'elle est stricte. Pour démontrer la deuxième inclusion, considérons dans l'anneau principal K la suite croissante d'idéaux $(d_1) \subset (d_2) \subset \ldots$ On vérifie directement que $D = \bigcup_i (d_i)$ est un idéal de K, donc $D = (d)$, $d \in D$. Par définition, $d \in (d_m)$ pour un certain m, d'où $(d_m) = (d_{m+1}) = \ldots$ La stabilisation de la suite croissante d'idéaux après un nombre fini de pas entraîne l'interruption de la suite de diviseurs non inversibles $d_1, d_2, d_3, \ldots$, avec $d_i \mid d_{i-1}$, et donc l'existence dans K d'une décomposition en éléments non factorisables. L'unicité de la décomposition dans K est conséquence des mêmes causes: $(a, b) \equiv aK + bK = dK = (d) \Rightarrow \Rightarrow d = \text{P.G.C.D.}\ (a, b) = ax + by$. Les raisonnements ultérieurs répètent la démonstration du théorème 3 (ii) du chapitre 5, § 3.

Les idéaux $(2, X)$ de $\mathbb{Z}[X]$ et (X, Y) de $\mathbb{R}[X, Y]$ ne sont pas principaux (voir chap. 5, § 2, n° 3, exemple). En même temps, les anneaux $\mathbb{Z}[X]$ et $\mathbb{R}[X, Y]$ sont factoriels d'après le théorème 2. On établit par là même la vérité de la suite (3). ■

Les anneaux principaux présentent de l'intérêt au point de vue purement algébrique, parce qu'ils se caractérisent par les propriétés des êtres naturels tels que les noyaux d'homomorphismes. D'autre part, les anneaux euclidiens sont plus commodes pour l'étude vu que dans ces anneaux l'algorithme de division euclidienne est valable.

2. Structures relatives à la théorie des anneaux. — Nous avons déjà à notre disposition un arsénal assez important de types d'anneaux et de moyens permettant de construire de nouveaux anneaux à partir de leur collection donnée. Comme exemples on peut signaler les constructions de l'anneau des matrices $M_n(K)$, du corps des quotients $Q(K)$ et de l'anneau des polynômes $K[X_1, \ldots, X_n]$, où K est un anneau commutatif (intègre dans le cas de $Q(K)$). Il est utile d'analyser encore, ne serait-ce que sommairement, les analogues qu'on trouve en théorie des anneaux, pour des faits géné-

raux concernant les homomorphismes, qui ont été établis pour les groupes au chapitre 7. Les démonstrations ne diffèrent généralement en rien de celles données dans le cas des groupes; elles seront laissées en exercice au lecteur.

Nous compléterons le théorème fondamental d'homomorphie des anneaux (chap. 4, § 4, théorème 2) de deux théorèmes sur leur isomorphie.

Théorème 3. — *Soient K un anneau, L un sous-anneau, J un idéal de K. Alors $L + J = \{x + y \mid x \in L, y \in J\}$ est un sous-anneau de K, contenant J comme idéal, $L \cap J$ est un idéal de L. L'application*

$$\varphi: x + J \mapsto x + L \cap J, \quad x \in L,$$

réalise l'isomorphisme des anneaux

$$(L + J)/J \cong L/L \cap J.$$

Démonstration. — Les deux premières assertions sont tout à fait évidentes. Pour prouver la dernière assertion, il convient de considérer la restriction $\pi_0 = \pi\,|_L$ de l'épimorphisme naturel π: $K \to K/J$. Son image Im π_0 est constituée des classes $x + J$, $x \in L$, c'est-à-dire Im $\pi_0 = (L + J)/J$. Le noyau Ker π_0 de l'épimorphisme π_0: $L \to (L + J)/J$ se compose des éléments $x \in L$ pour lesquels $x + J = J$. Par conséquent, Ker $\pi_0 = L \cap J$. D'après le théorème fondamental d'homomorphie, la relation $\bar{\pi}_0$: $x + L \cap J \mapsto \pi_0(x) = x + J$ établit l'isomorphisme: $L/L \cap J \cong (L + J)/J$. Il reste à remarquer que $\varphi = \bar{\pi}_0^{-1}$. ∎

Nous avons développé ce raisonnement, qui n'est qu'une copie de la démonstration du théorème 2 du chapitre 7, § 3, pour souligner une analogie complète avec la théorie des groupes.

Théorème 4. — *Soient K un anneau, J, L ses sous-anneaux, J étant un idéal de K et $J \subset L$. Alors, $\bar{L} = L/J$ est un sous-anneau de K/J et π^* : $L \to \bar{L}$ est une application bijective de l'ensemble $\Omega(K, J)$ des sous-anneaux de K, contenant J, sur l'ensemble $\Omega(\bar{K})$ de tous les sous-anneaux de l'anneau $\bar{K}$. Si $L \in \Omega(K, J)$, alors L est un idéal de K si, et seulement si, $\bar{L}$ est un idéal de $\bar{K}$. Ceci étant, on a*

$$K/L \cong \bar{K}/\bar{L} = (K/J)\,/\,(L/J).$$

Démonstration. — Elle représente un exercice facile (voir chap. 7, § 3, démonstration du théorème 3). ∎

Corollaire. — *Soit K un anneau commutatif ayant un élément unité 1. Un idéal J de K est maximal si, et seulement si, l'anneau quotient K/J est un corps.* ∎

Sur l'ensemble des idéaux de l'anneau K sont définies les opérations suivantes :

$$\textit{somme}: \quad J_1+J_2=\{x_1+x_2 \mid x_k\in J_k\},$$
$$\textit{intersection}: \quad J_1\cap J_2=\{x \mid x\in J_1,\ x\in J_2\},$$
$$\textit{produit}: \quad J_1J_2=\Big\{\sum_i x_{1i}x_{2i} \mid x_{ki}\in J_k\Big\}\subset J_1\cap J_2.$$

On peut aussi parler de la somme, de l'intersection et du produit de tout nombre fini d'idéaux. On a l'assertion suivante :

Proposition. — *Si les idéaux J, J_1, ..., J_n d'un anneau unitaire K vérifient les égalités*

$$J+J_k=K, \quad k=1, \ldots, n,$$

on a aussi les égalités

$$J+J_1\cap J_2\cap \ldots \cap J_n=K=J+J_1J_2 \ldots J_n.$$

Démonstration. — Comme $J_1J_2 \ldots J_n \subset J_1\cap J_2\cap \ldots \cap J_n$, il suffit d'établir l'égalité $J+J_1J_2 \ldots J_n=K$. Pour $n=1$, elle est vraie par hypothèse. Pour $n=2$, on a

$$1=1^2=(x_1+y_1)(x_2+y_2)=x+y_1y_2,$$

où x_1, x_2, $x\in J$, $y_i\in J_i$. Par suite, $1\in J+J_1J_2$ et $K=J+J_1J_2$. Plus loin, on a une récurrence évidente sur le nombre n. ■

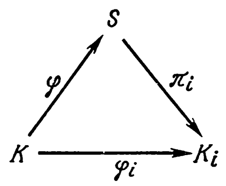
Soient K_1, ..., K_n une famille finie d'anneaux, $K=K_1\times \ldots \times K_n$ le produit cartésien des ensembles. Introduisons sur K une structure d'anneau en définissant les opérations d'addition et de multiplication de la façon suivante :

$$(x_1, \ldots, x_n)+(y_1, \ldots, y_n)=(x_1+y_1, \ldots, x_n+y_n);$$
$$(x_1, \ldots, x_n)\cdot(y_1, \ldots, y_n)=(x_1y_1, \ldots, x_ny_n).$$

Nous obtenons une *somme directe extérieure* $K=K_1\oplus \ldots \oplus K_n$ *des anneaux* K_i. Chacune des composantes K_i est une image par l'épimorphisme $\pi_i: K\to K_i$, $\pi_i: (x_1, \ldots, x_n)\mapsto x_i$, $1\leqslant i\leqslant n$. Si, ensuite, $J_i=\{(0, \ldots, x_i, \ldots, 0) \mid x_i\in K_i\}$, alors $J_i\cong K_i$, J_i est un idéal de K et $K=J_1+\ldots+J_n$.

Soit maintenant K un anneau contenant comme idéaux J_1, ..., J_n, tels que $K=J_1+\ldots+J_n$ et $J_k\cap\Big(\sum_{j\neq k} J_j\Big)=0$, $1\leqslant k\leqslant n$. Alors $K=J_1\oplus \ldots \oplus J_n$ est une *somme directe intérieure des idéaux* J_k. De même qu'en Théorie des groupes, la différence entre les sommes directes extérieure et intérieure des anneaux n'a qu'un sens purement ensembliste, de sorte qu'on peut utiliser la même désignation pour les deux sommes.

3. Applications à la théorie des nombres.— La propriété universelle des sommes directes est la suivante : *si* $S = K_1 \oplus \ldots \oplus K_n$ *et* K *est un anneau arbitraire à homomorphismes donnés* $\varphi_i : K \to K_i$, *il existe alors un seul homomorphisme* $\varphi = (\varphi_1, \ldots, \varphi_n) : K \to S$ *de noyau* $\operatorname{Ker} \varphi = \bigcap \operatorname{Ker} \varphi_i$, *qui rend commutatifs les diagrammes triangulaires*



pour $i = 1, \ldots, n$. Appliquons cette assertion évidente à un anneau K ayant les idéaux $J_1, \ldots, J_n$, et à la somme directe

$$S = K/J_1 \oplus \ldots \oplus K/J_n.$$

En posant $\varphi_i : K \to K/J_i = K_i$, on obtient un homomorphisme

$$\varphi : x \mapsto (x + J_1, \ldots, x + J_n) \tag{4}$$

de l'anneau K dans S, de noyau $\operatorname{Ker} \varphi = J_1 \cap \ldots \cap J_n$.

THÉORÈME 5 (théorème chinois sur les restes). — *Si, dans les conditions indiquées plus haut,* K *est un anneau unitaire et* $J_i + J_j = K$ *pour* $1 \leqslant i \neq j \leqslant n$, *l'application* φ (voir (4)) *est un épimorphisme.*

DÉMONSTRATION. — Il nous faut prouver que, quels que soient les éléments donnés $x_1, \ldots, x_n \in K$, il existe un $x \in K$ tel que $x_i + J_i = x + J_i$, c'est-à-dire $x - x_i \in J_i$, $i = 1, 2, \ldots, n$. Pour $n = 1$, ceci est évident, et pour $n = 2$ prenons des éléments $a_1 \in J_1$, $a_2 \in J_2$ pour lesquels $a_1 + a_2 = 1$, et posons $x = x_1a_2 + x_2a_1$. Alors

$$x - x_1 = (x_1a_2 + x_2a_1) - x_1(a_1 + a_2) = (x_2 - x_1)a_1 \in J_1,$$

$$x - x_2 = (x_1a_2 + x_2a_1) - x_2(a_1 + a_2) = (x_1 - x_2)a_2 \in J_2.$$

Raisonnons ensuite par récurrence sur n. Supposons trouvé un élément y tel que $y - x_i \in J_i$, $i = 1, 2, \ldots, n - 1$. Puisque par hypothèse, $J_i + J_n = K$, $1 \leqslant i \leqslant n - 1$, on a d'après la proposition du nº 2, $J_1 \cap \ldots \cap J_{n-1} + J_n = K$. En appliquant le cas de $n = 2$ que nous venons de considérer, aux idéaux $J_1 \cap \ldots \cap J_{n-1}$, J_n et aux éléments $y, x_n \in K$, on trouve $x \in K$ tel que $x - y \in J_1 \cap \ldots \cap J_{n-1}$, $x - x_n \in J_n$. Mais $x - y \in J_1 \cap \ldots \cap J_{n-1} \Rightarrow x - y \in J_i$, $1 \leqslant i \leqslant n - 1$. Compte tenu du choix

de y, on obtient

$$x - x_i = (x - y) + (y - x_i) \in J_i, \quad 1 \leqslant i \leqslant n - 1.$$

On voit que l'élément x satisfait à toutes les conditions requises. ■

Dans le théorème 5 et dans les raisonnements qui le précèdent, l'anneau K n'était pas supposé commutatif. Soient maintenant K un anneau intègre et $a_1, \ldots, a_n$ ses éléments premiers entre eux deux à deux, c'est-à-dire $a_iK + a_jK = K$ pour $i \neq j$ (dans un anneau factoriel K cette définition est compatible avec celle des éléments premiers entre eux, qui résulte de la décomposition de a_i en facteurs premiers). En écrivant l'appartenance $x - x_i \in a_iK$ sous la forme d'une congruence modulo l'idéal principal a_iK, nous utilisons comme à l'habitude la notation $x \equiv x_i \pmod{a_i}$.

COROLLAIRE 1. — *Soient K un anneau intègre et $a_1, \ldots, a_n$ ses éléments premiers entre eux deux à deux. Alors, quels que soient $x_1, \ldots$ $\ldots, x_n \in K$, il existe un élément $x \in K$ tel que $x \equiv x_i \pmod{a_i}$, $i = 1, \ldots, n$.* ■

COROLLAIRE 2. — *Soient n un entier naturel admettant la décomposition canonique $n = p_1^{m_1}, \ldots, p_r^{m_r}$; $Z_n = \mathbb{Z}/n\mathbb{Z}$ l'anneau de classes résiduelles modulo n, et $U(Z_n)$ le groupe multiplicatif de ses éléments inversibles. Alors*

(i) $Z_n \cong Z_{p^{m_1}} \oplus \ldots \oplus Z_{p^{m_r}}$ *(somme directe des anneaux)*;

(ii) $U(Z_n) \cong U(Z_{p^{m_1}}) \times \ldots \times U(Z_{p^{m_r}})$ *(produit direct des groupes).*

DÉMONSTRATION. — (i) En remplaçant dans (4) n par r et en posant $K = \mathbb{Z}$, $J_i = p_i^{m_i}\mathbb{Z}$ et $S = Z_{p_1^{m_1}} \oplus \ldots \oplus Z_{p_r^{m_r}}$, on obtient l'homomorphisme $\varphi: \mathbb{Z} \to S$ de noyau $\operatorname{Ker} \varphi = \bigcap J_i = n\mathbb{Z}$. L'épimorphie de φ se déduit du théorème 5, vu que P.G.C.D. $(p_i, p_j) = 1$ pour $i \neq j$.

(ii) Comme dans une somme directe arbitraire $K = K_1 \oplus \ldots$ $\ldots \oplus K_r$ les composantes K_i sont telles que $K_iK_j = 0$, $i \neq j$, la définition des éléments inversibles entraîne directement que $U(K) = U(K_1) \times \ldots \times U(K_r)$. Il ne reste qu'à appliquer ces résultats à la décomposition (i). ■

REMARQUE. — De l'assertion (ii) on déduit tout de suite que $\varphi(n) = \prod_{i=1}^{r} \varphi(p^{m_i})$, et comme $\varphi(p^m) = p^{m-1}(p - 1)$, on obtient de nouveau la formule pour les valeurs de la fonction d'Euler (voir § 1, nº 4, exemple 1). L'ordre de l'élément d'un groupe fini étant diviseur de l'ordre de ce groupe, on a

$$a^{\varphi(n)} \equiv 1 \pmod{n}$$

pour tout entier a premier avec n (généralisation du théorème de Fermat, connue sous le nom de théorème d'Euler).

Etant donné le corollaire 2, pour comprendre définitivement la structure du groupe $U(Z_n)$, il suffit de considérer le cas où $n = p^m$.

THÉORÈME 6. — *Soit m un entier positif.*

(i) *Si p est un nombre premier impair, alors $U(Z_{p^m})$ est un groupe cyclique.*

(ii) *Les groupes $U(Z_2)$ et $U(Z_4)$ sont des groupes cycliques d'ordres 1 et 2 respectivement, alors que $U(Z_{2^m})$, $m \geqslant 3$, est un produit direct d'un groupe cyclique d'ordre 2^{m-2} et d'un groupe cyclique d'ordre 2.*

DÉMONSTRATION.— (i) Un entier t premier avec n est, par définition, d'*ordre r modulo n*, si $|\langle t + n\mathbb{Z}\rangle| = r$, c'est-à-dire $t^r \equiv 1 \pmod{n}$ mais $t^k \not\equiv 1 \pmod{n}$ pour $0 < k < r$. Pour $r = \varphi(n)$ on dit que l'on a affaire à une *racine primitive t modulo n*. Généralement, on prend t dans le système des résidus $0, 1, \ldots, n - 1$ modulo n, mais nous ne fixons aucun système.

D'après le théorème 5 du § 1, le groupe $Z_p^* = U(Z_p)$ est cyclique, c'est-à-dire il existe une racine primitive a_0 modulo p. Comme $a_0^{p^{m-1}} \equiv a_0 \pmod{p}$, l'entier $a = a_0^{p^{m-1}}$ sera racine primitive modulo p. D'autre part, $a^{p-1} = a_0^{p^{m-1}(p-1)} = a_0^{\varphi(p^m)} \equiv 1 \pmod{p^m}$. Par conséquent, la classe $\bar{a} = a + p^m\mathbb{Z}$ engendre dans $U(Z_{p^m})$ un sous-groupe cyclique d'ordre $p - 1$.

On a ensuite

$$(1+p)^p = \sum_{i=0}^{p} \binom{p}{i} p^i = 1 + p^2 + \frac{1}{2}(p-1)p^3 + \sum_{i\geqslant 3} \binom{p}{i} p^i.$$

Comme $p > 2$, on a $(1+p)^p \equiv 1 + p^2 \pmod{p^3}$. En supposant par récurrence que $(1+p)^{p^j} \equiv 1 + p^{j+1} \pmod{p^{j+2}}$, on obtient

$$(1+p)^{p^{j+1}} = [1 + (1+sp)p^{j+1}]^p = \sum_{i=0}^{p} \binom{p}{i} (1+sp)^i p^{(j+1)i} =$$

$$= 1 + (1+sp)p^{j+2} + \frac{1}{2}(p-1)(1+sp)^2 p^{2(j+1)+1} + \ldots$$

d'où $(1+p)^{p^{j+1}} \equiv 1 + p^{j+2} \pmod{p^{j+3}}$. En particulier, $(1+p)^{p^{m-1}} \equiv 1 \pmod{p^m}$. Or, $(1+p)^{p^{m-2}} \equiv 1 + p^{m-1} \not\equiv 1 \pmod{p^m}$ et donc, la classe $\bar{b} = 1 + p + p^m\mathbb{Z}$ ayant pour représentant $b = 1 + p$ engendre dans $U(Z_{p^m})$ un groupe cyclique d'ordre p^{m-1}. Suivant la proposition du chapitre 4, § 2, nº 3, les éléments $\bar{a}$, $\bar{b}$ d'ordres $p - 1$, p^{m-1} premiers entre eux, engendrent le groupe cyclique $\langle \bar{a}\bar{b}\rangle$ d'ordre $p^{m-1}(p-1) = \varphi(p^m) = |U(Z_{p^m})|$.

(ii) En ce qui concerne les groupes $U(Z_2)$ et $U(Z_4)$, tout est clair. Pour $m > 2$, en partant de la congruence triviale $5 \equiv 1 + 2^2 \pmod{2^3}$, on vérifie facilement par récurrence sur j que

$$5^{2^j} \equiv 1 + 2^{j+2} \pmod{2^{j+3}}.$$

En particulier,

$$5^{2^{m-3}} \equiv 1 + 2^{m-1} \not\equiv 1 \pmod{2^m}, \quad 5^{2^{m-2}} \equiv 1 \pmod{2^m},$$

de sorte que 5 est d'ordre 2^{m-2} modulo 2^m, et la classe $5 + 2^m\mathbb{Z}$ engendre dans $U(Z_{2^m})$ un sous-groupe cyclique d'indice 2. Remarquons que $-1 + 2^m\mathbb{Z} \notin \langle 5 + 2^m\mathbb{Z}\rangle$, puisque $5^j \equiv -1 \pmod{2^m} \Rightarrow 5^j \equiv -1 \pmod{4} \Rightarrow 1 \equiv -1 \pmod{4}$ est une contradiction. Comme $|\langle -1 + 2^m\mathbb{Z}\rangle| = 2$,

$$U(\mathbb{Z}/2^m\mathbb{Z}) = \langle 5 + 2^m\mathbb{Z}\rangle \times \langle -1 + 2^m\mathbb{Z}\rangle$$

est un 2-groupe abélien de type $(2^{m-2}, 2)$ (voir chap. 7, § 5). ∎

COROLLAIRE. — *Le groupe $U(Z_n)$ est cyclique (ou ce qui revient au même, la racine primitive modulo n existe) si, et seulement si, l'entier $n > 1$ est de la forme 2, 4, p^m ou $2p^m$, où p est un nombre premier impair.* ∎

EXERCICES

1. Démontrer qu'un élément non nul p d'un anneau factoriel K est premier si, et seulement si K/pK est un anneau intègre.

2. Démontrer que, si un anneau intègre K n'est pas un corps, alors $K[X]$ n'est pas un anneau principal.

3. Montrer que les éléments $x + y\sqrt{-3}$, avec $x, y \in \mathbb{Z}$ ou encore $x = \frac{2k+1}{2}$, $y = \frac{2l+1}{2}$, $k, l \in \mathbb{Z}$, forment un anneau intègre K. Vérifier que c'est un anneau euclidien avec la fonction $\delta = N$ (norme dans $\mathbb{Q}(\sqrt{-3})$). Montrer que le sous-anneau $\mathbb{Z}[\sqrt{-3}] \subset K$ n'est pas même factoriel.

4. Trouver tous les éléments premiers de l'anneau des entiers de Gauss.

5. Perfectionner le corollaire 1 du théorème 5 dans le cas d'un anneau factoriel K, en introduisant à cet effet, en plus des éléments $a_1, \ldots, a_n$ premiers entre eux deux à deux, encore les éléments $\tilde{a}_i = \prod_{j \neq i} a_j$. Trouver $b_i \in K$ pour lesquels $b_i \equiv 1 \pmod{a_i}$, $b_i \equiv 0 \pmod{\tilde{a}_i}$, $1 \leqslant i \leqslant n$. Soient $x_1, \ldots, x_n \in K$. Introduire un élément $x = \sum b_i x_i$ et vérifier que $x \equiv x_i \pmod{a_i}$, $1 \leqslant i \leqslant n$ (ce qui est bien commode surtout dans le cas où l'on a affaire à un grand nombre de collections $x_1, \ldots, x_n$).

6. Appliquer l'exercice précédent aux modules $a_1 = 5$, $a_2 = 9$ et aux couples $(x_1, x_2) = (2, 5), (3, 2), (3, 5)$. Que peut-on dire de l'ordre de x modulo 45 ?

7. Soit p un nombre premier impair. Si la congruence $x^2 \equiv a \pmod{p}$ admet une solution, l'entier a s'appelle *résidu quadratique modulo p*, dans le cas contraire il s'appelle *non-résidu quadratique*. Le *symbole de Legendre* $\left(\frac{a}{p}\right)$ est défini par la relation

$$\left(\frac{a}{p}\right) = \begin{cases} 0 \text{ si } a \equiv 0 \pmod{p}, \\ 1 \text{ si } a \not\equiv 0 \pmod{p} \text{ est un résidu quadratique}, \\ -1 \text{ si } a \not\equiv 0 \pmod{p} \text{ est un non-résidu quadratique}. \end{cases}$$

Montrer que $\left(\frac{a}{p}\right)=1 \Leftrightarrow a+p\mathbb{Z}\in\mathbb{Z}_p^{*2}$ et $\left(\frac{a}{p}\right)\equiv a^{\frac{p-1}{2}}$ (mod. p). Puis, vérifier que $\left(\frac{ab}{p}\right)=\left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$ et que le nombre de résidus quadratiques dans le système $1, 2, \ldots, p-1$ coïncide avec le nombre de non-résidus. Vérifier pour de petits nombres premiers impairs p et q la *loi de réciprocité des résidus quadratiques*

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right)=(-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}},$$

démontrée (par plusieurs procédés) dans le cas général par Gauss. Déduire de l'exemple 1 du n° 1 la relation $\left(\frac{-1}{p}\right)=(-1)^{\frac{p-1}{2}}$.

8. Démontrer (en utilisant les notations de l'exercice précédent) que $\left(\frac{2}{p}\right)= (-1)^{\frac{p^2-1}{8}}$, c'est-à-dire que 2 est un carré mod. p si, et seulement si, $p\equiv\pm 1$ (mod. 8). (I n d i c a t i o n. En s'appuyant sur l'exercice 9 du § 1, considérer la racine primitive 8-ième α de l'unité dans la clôture algébrique Ω_p du corps commutatif $\mathbb{F}_p$. Comme $\alpha^4=-1$, on a $\alpha^2+\alpha^{-2}=0$; en outre $\alpha^5=-\alpha$, $\alpha^{-5}=-\alpha^{-1}$, d'où $\alpha^5+\alpha^{-5}=-(\alpha+\alpha^{-1})$. En posant $\beta=\alpha+\alpha^{-1}$, on a $\beta^2=\alpha^2+\alpha^{-2}+2=2$, de sorte que $p\equiv\pm 1$ (mod. 8) $\Rightarrow \beta^p=\alpha^p+ +\alpha^{-p}=\alpha+\alpha^{-1}=\beta\Rightarrow 1=\beta^{p-1}=(\beta^2)^{\frac{p-1}{2}}=2^{\frac{p-1}{2}}\Rightarrow\left(\frac{2}{p}\right)=1$. De façon analogue, $\beta\equiv\pm 5$ (mod. 8) $\Rightarrow \beta^p=\alpha^p+\alpha^{-p}=\alpha^5+\alpha^{-5}=-(\alpha+\alpha^{-1})= =-\beta\Rightarrow -1=\beta^{p-1}=2^{\frac{p-1}{2}}\Rightarrow\left(\frac{2}{p}\right)=-1$.)

9. (complément du chap. 3, § 2, n° 5). Soit $f(X)=f(\ldots, x_{ij}, \ldots)$ un polynôme non nul à coefficients dans $\mathbb{Z}$ ou dans un corps commutatif, de n^2 indéterminées indépendantes $x_{ij}\in K$, $1\leqslant i, j\leqslant n$, considéré comme fonction de la matrice $X=(x_{ij})$. Démontrer que, si $f(XY)=f(X)f(Y)$ pour tous les $X, Y\in M_n(K)$, alors $f(X)=(\det X)^m$, où m est un entier non négatif. En particulier, $f(X)=\det X$ si $f(\operatorname{diag}(x, 1, \ldots, 1))=x$. (I n d i c a t i o n. Si $n=1$ et $f(x)=\sum_{i=1}^{m} a_i x^i$, $a_m\neq 0$, on a

$$f(xy)=\sum_{i=1}^{m} a_i x^i y^i=f(x)f(y)=f(x)\Big(\sum_{i=1}^{m} a_i y^i\Big),$$

où x, y sont des indéterminées indépendantes. L'égalité des coefficients de y^m montre que $a_m x^m=f(x)a_m$ et donc $f(x)=x^m$. Si maintenant on pose dans le cas général $g(x)=f(x\cdot E)$, alors $g(xy)=g(x)g(y)$. Ce resultat joint au fait que l'assertion est vraie pour $n=1$, entraîne que $g(x)=x^s$. Comme $X\cdot X^{\vee}= =(\det X)\cdot E$, on a

$$f(X)f(X^{\vee})=f((\det X)E)=g(\det X)=(\det X)^s.$$

Or, $f(X)$, $f(X^{\vee})$ et $\det X$ sont des polynômes en x_{ij}, $1\leqslant i, j\leqslant n$, et $\det X$ est irréductible (voir chap. 5, § 3, exercice 7). D'après le théorème 2, selon lequel l'anneau des polynômes d'un nombre quelconque d'indéterminées est factoriel, il vient que $f(X)=c(\det X)^m$, c étant une constante, et $f(XY)=f(X)f(Y)\Rightarrow \Rightarrow c^2=c$. Comme $c\neq 0$, on a $c=1$.)

10. Montrer que l'anneau $Q_M(\mathbb{Z})$ de tous les nombres rationnels a/b, avec b

non divisible par un nombre premier fixe p (voir chap. 5, § 4, exercice 6), contient un idéal maximal et un seul

$$J = \{a/b \in Q_M(\mathbb{Z}) \mid p \text{ divise } a\}.$$

Tout anneau qui contient un seul idéal maximal s'appelle *anneau local*. (Indication. Il est évident que J est un idéal propre de $Q_M(\mathbb{Z})$). Si $c/d \notin J$, alors $c \notin p\mathbb{Z}$, et par conséquent, $d/c \in Q_M(\mathbb{Z})$. Ceci signifie que tout idéal L obtenu de J par adjonction d'au moins un élément c/d contient $1 = c/d \odot d/c$ et donc coïncide avec $Q_M(\mathbb{Z})$.

11. Montrer que dans tout anneau local K contenant $\mathfrak{m}$ comme idéal maximal, les éléments n'appartenant pas à $\mathfrak{m}$ sont inversibles.

12. Un idéal $\mathfrak{p}$ d'un anneau unitaire K s'appelle idéal premier si l'anneau quotient $K/\mathfrak{p}$ est intègre. Tout idéal maximal est premier. Le complémentaire $M = K \setminus \mathfrak{p}$ de $\mathfrak{p}$ dans l'anneau K est un sous-ensemble multiplicatif (un monoïde ne contenant pas 0). Ceci étant, l'anneau $Q_M(K)$ est désigné le plus souvent par le symbole $M^{-1}K$ ou tout simplement $K_\mathfrak{p}$. Montrer que l'anneau $K_\mathfrak{p}$ est toujours local et que son idéal maximal $\mathfrak{m}_\mathfrak{p}$ se compose de quotients de la forme a/b, où $a \in \mathfrak{p}$ et $b \in K \setminus \mathfrak{p}$. Montrer aussi que $\mathfrak{m}_\mathfrak{p} \cap K = \mathfrak{p}$.

L'opération de passage de K à l'anneau local $K_\mathfrak{p}$ s'appelle localisation de l'anneau K par rapport à l'idéal premier $\mathfrak{p}$.

§ 3. Modules

La notion de module est porteur d'un principe fondamental élaboré en algèbre il y a un demi-siècle. Ce principe consiste en ce que l'étude de tout système algébrique doit porter non seulement sur les propriétés internes de ce système mais aussi sur toutes ses représentations (dans toute la force du terme).

1. Généralités sur les modules.— Commençons par donner la définition classique d'un module. Soient K un anneau associatif unitaire et V un groupe abélien noté additivement. Soit ensuite donnée une application $(x, v) \mapsto xv$ de $K \times V$ dans V, satisfaisant aux conditions

(M1) $\quad x(u + v) = xu + xv,$

(M2) $\quad (x + y)v = xv + yv,$

(M3) $\quad (xy)v = x(yv),$

(M4) $\quad 1 \cdot v = v$

quels que soient $x, y \in K$, $u, v \in V$. Alors V s'appelle *K-module à gauche* (ou *module à gauche sur l'anneau K*). On définit de façon analogue un K-module à droite. Dans ce qui suit nous parlons tout simplement d'un K-module, bien que dans certaines situations les deux genres de modules apparaissent ensemble.

L'axiome (M4) (condition pour que le module soit unitaire) est naturellement superflu si l'anneau K n'a pas d'élément unité. Un fait plus important est que l'axiome (M3) admet des modifications adaptées à certains anneaux non associatifs. Un exemple de module sur un anneau non associatif est fourni à la fin du paragraphe sui-

vant. Pour le moment, nous allons partir de la définition donnée plus haut.

Soit V un K-module. On dit qu'un sous-groupe $U \subset V$ est un *sous-module* de V, si $xu \in U$ pour tous les $x \in K$, $u \in U$.

Soient ensuite U et V deux K-modules arbitraires. On appelle *homomorphisme des K-modules* (ou tout simplement *K-homomorphisme*) de U dans V une application $\sigma: U \to V$ telle que

$$\sigma(u_1 + u_2) = \sigma(u_1) + \sigma(u_2),$$

$$\sigma(xu) = x\sigma(u)$$

quels que soient u_1, u_2, $u \in U$, $x \in K$. On vérifie sans peine que $\operatorname{Ker} \sigma = \{u \in U \mid \sigma(u) = 0\}$ est un K-sous-module de U et que l'image $\operatorname{Im} \sigma$ est un K-sous-module de V.

A tout sous-module $U \subset V$ sur K est associé un *module quotient* $V/U = \{v + U \mid v \in V\}$ qui est un groupe quotient du groupe abélien additif, avec l'opération de K définie par

$$x(v + U) = xv + U.$$

Le théorème fondamental d'homomorphie et les deux théorèmes d'isomorphie que nous avons démontrés pour les groupes (chap. 7, § 3) et ensuite pour les anneaux, s'appliquent mot pour mot au cas des modules avec de petites modifications apportées dans leurs démonstrations.

Après le § 2 du chapitre 7, où nous avons considéré les axiomes de type (M3), (M4), et après le chapitre 8 bien substantiel, consacré aux représentations des groupes (axiomes (M1), (M3), (M4)), il est peu probable que les exemples de K-modules que nous allons fournir, produisent l'impression qu'il s'agit d'une nouveauté. Néanmoins, il est utile de les discuter et de les comparer.

1) Tout groupe abélien A est un $\mathbb{Z}$-module. En effet, l'application $(n, a) \mapsto na$ de $\mathbb{Z} \times A$ dans A satisfait à tous les axiomes (M1) à (M4). Cette conception des groupes abéliens comme modules sur $\mathbb{Z}$ se montre très utile.

2) Tout groupe abélien A est un module sur son *anneau d'endomorphismes* $\operatorname{End} A$. Par définition, $\operatorname{End} A$ se compose de toutes les applications $\varphi: A \to A$ satisfaisant à la condition $\varphi(a + a') = \varphi(a) + \varphi(a')$. Les opérations d'addition et de multiplication dans $\operatorname{End} A$ sont introduites de façon naturelle: $(\varphi + \psi)(a) = \varphi(a) + \psi(a)$, $(\varphi\psi)(a) = \varphi(\psi(a))$; $1(x) = x$, $0(x) = 0$. L'application $(\varphi, a) \mapsto \varphi(a)$ de $\operatorname{End} A \times A$ dans A confère évidemment à A une structure de $\operatorname{End} A$-module.

3) Un espace vectoriel V sur un corps commutatif P est sûrement un P-module. Si nous avons encore un opérateur linéaire donné $\mathcal{A}: V \to V$, nous pouvons conférer à V une structure de module $V_{\mathcal{A}}$ sur l'anneau des polynômes $P[X]$, en posant

$$f(X)\,v = f(\mathcal{A})\,v = \alpha_0 v + \alpha_1 \mathcal{A} v + \ldots + \alpha_k \mathcal{A}^k v$$

pour toute $v \in V$ et tout polynôme $f \in P[X]$. Les axiomes (M1) à (M4) sont vérifiés puisque l'opérateur $f(A)$ est, de même que $\mathcal{A}$, linéaire et

$$(f+g)(\mathcal{A}) = f(\mathcal{A}) + g(\mathcal{A}), \quad (fg)(\mathcal{A}) = f(\mathcal{A})\, g(\mathcal{A})$$

(propriété universelle des anneaux des polynômes; voir chap. 5, § 2). Les sous-modules de $V_{\mathcal{A}}$ seront les sous-espaces $\mathcal{A}$-invariants. Aux opérateurs linéaires différents du même espace V correspondent en général des $P[X]$-modules différents (non isomorphes).

4) Un idéal à gauche arbitraire J d'un anneau K (c'est-à-dire, un sous-groupe additif pour lequel $a \in J \Rightarrow ba \in J$; $\forall b \in K$) est muni d'une structure naturelle de K-module avec l'opération $(x, y) \mapsto xy$, $x \in K$, $y \in J$, induite par l'opération de multiplication dans K. Dans le cas où $J = K$, l'anneau K est considéré comme un module ${}_K K$ sur lui-même. Un tel concept de K aboutit à des résultats bien féconds.

5) En revenant à l'exemple précédent, construisons le module quotient $K/J = \{y + J \mid y \in K\}$. Conformément à la définition générale, $(x, y + J) \mapsto xy + J$ est une opération de K sur K/J. Remarquons qu'étant un homomorphisme des K-modules, l'épimorphisme canonique $\pi: K \to K/J$ vérifie la relation $\pi(xy) = xy + J = x(y + J) = x\pi(y)$. Si J est un idéal bilatère, alors K/J est un anneau et π est un homomorphisme des anneaux: $\pi(xy) = \pi(x)\,\pi(y)$.

L'*intersection* $\bigcap_i V_i$ de toute famille de *sous-modules* $V_i \subset V$ sur K est un sous-module de V. En particulier, l'intersection de tous les sous-modules contenant un ensemble donné $T \subset V$ conduit à un *sous-module* $\langle T \rangle$ *engendré par l'ensemble* T et composé de tous les éléments de la forme $x_1t_1 + \ldots + x_kt_k$, où $x_i \in K$, $t_i \in T$. Remarquons à propos que les éléments non nuls $t_1, \ldots, t_k \in V$ sont dits *linéairement dépendants sur* K si $x_1t_1 + \ldots + x_kt_k = 0$, où tous les x_i ne sont pas simultanément nuls. Le sous-module engendré par la famille $\{V_1, \ldots, V_m\}$ de sous-modules V_i s'appelle leur *somme* et se note par le symbole usuel: $\sum V_i = V_1 + \ldots + V_m$.

Le module V sur K engendré par un seul élément v est dit *cyclique*. Il est de la forme $V = Kv = \{xv \mid x \in K\}$, où $v \in V$, et represente un analogue de groupe cyclique. En particulier, le module cyclique ${}_K K = K \cdot 1$ (voir exemple 4) est un analogue du groupe $(\mathbb{Z}, +)$.

Si $V = Kv_1 + \ldots + Kv_n$ est la somme d'un nombre fini de modules cycliques, on dit que le module V est un K-module *de type fini*.

On vérifie aisément que l'application $x \mapsto xv$ est un homomorphisme des modules ${}_K K \to Kv$. Son noyau $\mathrm{Ann}(v) = \mathrm{Ann}_K(v) = \{x \in K \mid xv = 0\}$ est un idéal à gauche de K, appelé *annulateur*

(ou *torsion*) de l'élément v. Ainsi, $Kv \cong K/\mathrm{Ann}\,(v)$. Un élément $v \in V$, avec $\mathrm{Ann}\,(v) \neq 0$, est dit *périodique*. Un module dont tous les éléments sont périodiques est dit, lui aussi, *périodique*. Si V ne contient pas d'éléments périodiques non nuls, on dit que le *module V est sans torsion*.

On appelle *annulateur* (ou *torsion*) d'un K-module V l'ensemble

$$\mathrm{Ann}\,(V) = \{a \in K \mid aV = 0\} = \bigcap_{v \in V} \mathrm{Ann}\,(v).$$

On dit qu'un module V est *exact* si $\mathrm{Ann}\,(V) = 0$.

On peut parvenir aux mêmes notions en empruntant un autre cheminement. Soit $V\,(x)$ un ensemble des éléments $v \in V$ annulés par l'élément $x \in K$. Si K est un anneau intègre, alors $V\,(x) + V\,(y) \subset V\,(xy)$, et on peut introduire la notion de *sous-module de torsion* $\mathrm{Tor}\,(V) = \sum_{x \in K} V\,(x)$. Dans le cas où $\mathrm{Tor}\,(V) = V$, on dit que V est un *module de torsion*. Lorsque $\mathrm{Tor}\,(V) = 0$, nous retrouvons la notion de module sans torsion.

Comme exemples caractéristiques de modules périodiques on peut indiquer: a) tout groupe abélien fini (module périodique de type fini sur $\mathbb{Z}$; la torsion est $m\mathbb{Z}$ si m est l'exposant du groupe); b) le module $V_{\mathcal{A}}$ sur $P\,[X]$ associé à l'opérateur linéaire $\mathcal{A}$ (voir exemple 3; la torsion est l'idéal principal engendré par le polynôme minimal de l'opérateur $\mathcal{A}$).

PROPOSITION 1.— $\mathrm{Ann}\,(V)$ *est toujours un idéal bilatère de l'anneau K. En posant* $(x + \mathrm{Ann}\,(V))\,v = xv$, *on confère à V une structure de* $(K/\mathrm{Ann}\,(V))$*-module exact.*

DÉMONSTRATION. — Posons $A = \mathrm{Ann}\,(V)$. Il est clair que A est un sous-groupe additif de K. On a $(xax')\,v = xa\,(x'v) = (xa)\,v' = x\,(av') = x \cdot 0 = 0$, quels que soient $x,\ x' \in K,\ a \in A,\ v \in V$, d'où il résulte justement que $KAK \subset A$, c'est-à-dire que A est un idéal bilatère de K. Si maintenant $x + A = x' + A$, alors $x - x' \in A$, d'où $(x - x')\,v = 0$ ou $xv = x'v$. Par suite, $(x + A)\,v = (x' + A)\,v$, c'est-à-dire l'opération de l'anneau quotient K/A sur V est définie correctement. Il n'est pas difficile de vérifier que V est un K/A-module pour cette opération. Enfin, on a

$$(x + A)\,V = 0 \Rightarrow x + A \in \mathrm{Ann}_{K/A}\,(V) \Rightarrow xV = 0 \Rightarrow x \in A.$$

Par conséquent, seul l'élément nul de K/A annule V. ■

De la proposition 1 il résulte que l'anneau quotient $K/\mathrm{Ann}\,(V)$ est isomorphe à un sous-anneau de l'anneau $\mathrm{End}\,(V)$ (voir exemple 2).

Si V, W sont deux K-modules, l'ensemble $\mathrm{Hom}_K\,(V,\ W)$ de tous les homomorphismes K-linéaires $\sigma\colon V \to W$ est un groupe

abélien pour l'opération d'addition des homomorphismes:

$$(\sigma + \tau)(xv) = \sigma(xv) + \tau(xv) = x\sigma(v) + x\tau(v) = \\ = x(\sigma(v) + \tau(v)) = x((\sigma + \tau)(v)).$$

Pour les modules V, W sur un anneau commutatif K, l'ensemble $\mathrm{Hom}_K(V, W)$ est lui-même un K-module si par $x\sigma$, $x \in K$, $\sigma \in \mathrm{Hom}_K(V, W)$, on entend l'application $v \mapsto x(\sigma(v))$:

$$(x\sigma)(yv) = x\cdot\sigma(yv) = x(y\sigma(v)) = (xy)(\sigma(v)) = \\ = (yx)(\sigma(v)) = y(x\sigma(v)) = y((x\sigma)(v)).$$

Dans le cas où $W = V$, l'ensemble $\mathrm{End}_K(V) = \mathrm{Hom}_K(V, V)$ est un anneau; la multiplication est représentée par la composition naturelle des K-homomorphismes $\varphi \circ \psi$: $(\varphi \circ \psi)(xv) = \varphi(\psi(xv)) = \varphi(x\psi(v)) = x\varphi(\psi(v)) = x((\varphi \circ \psi)(v))$. On ne perdra pas de vue qu'en considérant V comme groupe abélien additif, nous écrivons $\mathrm{End}_{\mathbb{Z}}(V)$ et qu'en général, $\mathrm{End}_K(V)$ est un sous-anneau propre de $\mathrm{End}_{\mathbb{Z}}(V)$. Dans le cas d'un espace vectoriel V sur un corps commutatif K, on écrit généralement $\mathcal{L}(V) = \mathrm{End}_K(V)$ et on l'appelle *anneau* (ou *algèbre*) *des opérateurs linéaires.*

L'anneau $\mathrm{End}_K(V)$ des K-endomorphismes du module V est encore appelé *centralisateur de l'anneau K sur V.* Son rôle est surtout important dans le cas des *modules irréductibles* (ou *premiers*). On dit qu'un module V sur un anneau K est irréductible, si: a) $V \neq 0$; b) 0, V sont les seuls sous-modules de V, c) $KV \neq 0$ (cette condition est automatiquement satisfaite si K a un élément unité). Il est clair qu'un *K-module $V \neq 0$ est irréductible si, et seulement si, $V = Kv$ quel que soit $v \neq 0$ de V.*

PROPOSITION 2 (lemme de Schur).— *Si V, W sont deux K-modules irréductibles et si σ est un K-homomorphisme non nul de V dans W, alors σ est un isomorphisme et* $\mathrm{End}_K(V)$ *est un corps pour tout K-module irréductible V.*

DÉMONSTRATION. — Elle est donnée au chapitre 8, § 4, où le même lemme de Schur (théorème 1) a été démontré pour les G-espaces irréductibles. ∎

2. Modules libres.— Nous dirons qu'un K-module V est une *somme directe (intérieure) de ses sous-modules* $V_1, \ldots, V_n$ si $V = V_1 + \ldots + V_n$ et $V_i \cap \sum_{j \neq i} V_j = 0$ pour $i = 1, \ldots, n$. En d'autres termes, $V = V_1 \oplus \ldots \oplus V_n$ (désignation de la somme directe de sous-modules) si tout élément $v \in V$ s'écrit de manière unique sous la forme d'une combinaison linéaire $v = v_1 + \ldots + v_n$, $v_i \in V_i$. Une *somme directe extérieure de K-modules* $V_1, \ldots, V_n$ est définie de façon évidente (de même que dans le cas des anneaux),

avec l'opération $x(v_1, \ldots, v_n) = (xv_1, \ldots, xv_n)$ de l'élément $x \in K$ sur la ligne $(v_1, \ldots, v_n)$, $v_i \in V_i$.

Soient ensuite V un K-module et $\{v_1, \ldots, v_n\}$ un sous-ensemble fini de V. On dit que $\{v_1, \ldots, v_n\}$ *engendre V librement*, si $V = Kv_1 + \ldots + Kv_n$ et toute application φ de l'ensemble $\{v_1, \ldots, v_n\}$ dans un K-module quelconque W se prolonge en un K-homomorphisme $\tilde{\varphi}: V \to W$, de manière que $\tilde{\varphi}(v_i) = \varphi(v_i)$, $1 \leqslant i \leqslant n$.

Un module V sur K, librement engendré par un sous-ensemble $\{v_1, \ldots, v_n\}$ s'appelle *module libre de rang m*, et $\{v_1, \ldots, v_n\}$ s'appelle sa *base* (*libre*) *sur K*.

PROPOSITION 3. *Les assertions suivantes sont équivalentes:*

(i) *l'ensemble* $\{v_1, \ldots, v_n\}$ *engendre V librement;*

(ii) *l'ensemble* $\{v_1, \ldots, v_n\}$ *est linéairement indépendant et* $\langle v_1, \ldots, v_n \rangle = V$;

(iii) *tout élément* $v \in V$ *s'écrit d'une manière et d'une seule sous la forme* $v = \sum x_i v_i$, $x_i \in K$;

(iv) $V = Kv_1 \oplus \ldots \oplus Kv_n$ *est une somme directe et* $\operatorname{Ann}(v_i) = 0$;

(v) $V \cong_K K \oplus \ldots \oplus_K K$ *est une somme directe de n termes égaux à* $_K K$ *(ainsi, un K-module libre de rang n par rapport à la base* $\{v_1, \ldots, v_n\}$ *est isomorphe au module* K^n *des lignes* $(x_1, \ldots, x_n)$ *de longueur n et de composantes* $x_i \in K$).

DEMONSTRATION. — Elle est analogue aux raisonnements développés au chapitre 2 pour les espaces vectoriels sur un corps commutatif, mais il convient de faire preuve d'une certaine prudence liée soit à la non-commutativité de l'anneau K, soit à l'existence d'éléments non inversibles dans K. ■

Il existe des exemples d'anneaux non commutatifs assez complexes, avec $K^m \cong K^n$ pour $m \neq n$, mais le comportement des anneaux commutatifs est bon sous ce rapport.

PROPOSITION 4. — *Le rang d'un module libre de type fini sur un anneau intègre K est défini univoquement.*

DEMONSTRATION. — Soient $\{v_1, \ldots, v_n\}$, $\{u_1, \ldots, u_m\}$ deux bases d'un module libre V sur K. Alors

$$v_j = \sum_{i=1}^{m} a_{ij} u_i, \quad u_i = \sum_{k=1}^{n} b_{ki} v_k.$$

On obtient pour les matrices $A = (a_{ij})$ et $B = (b_{kl})$ de types respectifs (m, n) et (n, m) les relations suivantes

$$AB = E_m, \quad BA = E_n.$$

27—0877

En plongeant K dans le corps des quotients $Q(K)$, nous obtiendrons à l'aide du théorème 4 du chapitre 2, § 4 (qui est vrai pour tout corps commutatif et non seulement pour $\mathbb{R}$) que $\min(n, m) \geqslant m$, $\min(n, m) \geqslant n$, d'où $m = n$. Ajoutons que le cas où $m < \infty$, $n = \infty$ est impossible à réaliser; parce que les expressions donnant u_i ne comportent qu'un nombre fini d'éléments de base v_k qui engendrent librement le module V tout entier. ■

Remarque— Dans le cas d'un anneau commutatif unitaire arbitraire K, on obtient le même résultat si l'on choisit dans K un idéal maximal J et l'on passe au corps commutatif K/J. Nous n'entrerons pas ici dans le détail de ce problème.

Signalons qu'à la différence de la situation qui se présente dans les espaces vectoriels, un ensemble arbitraire engendrant un K-module libre ne doit pas contenir nécessairement une base de ce module. Par exemple, deux nombres premiers différents p, q engendrent toujours ${}_{\mathbb{Z}}\mathbb{Z}$ car $up + vq = 1$ pour certains $u, v \in \mathbb{Z}$. Mais $\{p, q\}$ n'est pas une base, car $p \cdot q - q \cdot p = 0$, et $\mathbb{Z}p$, $\mathbb{Z}q$ sont des sous-modules propres de ${}_{\mathbb{Z}}\mathbb{Z}$.

Le rôle des modules libres est déterminé par leur définition.

Théorème 1.— *Tout K-module de type fini est une image homomorphe d'un K-module libre de type fini.*

Démonstration.—Soit $U = \sum_{i=1}^{n} Ku_i$ un K-module engendré par n éléments $u_1, \ldots, u_n$. Considérons un K-module libre V ayant pour base $\{v_1, \ldots, v_n\}$. Son existence est assurée par la proposition 3 (v). En vertu de la définition même du module libre, l'application $\varphi: v_i \mapsto u_i$ peut être prolongée en un K-homomorphisme $\tilde{\varphi}: V \to U$. L'image $\operatorname{Im} \tilde{\varphi}$ contient l'ensemble engendrant le module U et donc le module U tout entier. ■

Un sous-module d'un module libre n'est pas toujours libre, même s'il est l'un de ses termes directs. Voilà un exemple bien simple. Soient $K = Z_6$, $U = K(2 + 6\mathbb{Z})$, $V = K(3 + 6\mathbb{Z})$. Alors $K = U \oplus V$ est somme directe des K-modules U, V dont aucun n'est libre: $|K| = 6$, alors que $|U| = 3$, $|V| = 2$.

Théorème 2.—*Soit $V = Kv_1 \oplus \ldots \oplus Kv_n$ un module libre de rang n sur un anneau principal K. Alors, chacun de ses sous-modules U est un sous-module libre de rang $m \leqslant n$.*

Démonstration.—Soit d'abord $n = 1$, c'est-à-dire $V \cong K$. Tout sous-module $U \subset V$ est isomorphe à un idéal de K, et donc $U \cong (u) = Ku$. Lorsque $u = 0$, on a $U = 0$ (le sous-module nul peut être considéré comme un module libre de rang zéro). Si $u \neq 0$, alors $au \neq 0$ pour tout $0 \neq a \in K$, puisque K est un anneau intègre.

Par conséquent, U est un module libre (cyclique) de rang 1. Pour $n > 1$, raisonnons par récurrence. Considérons dans V le sous-module libre $V' = Kv_2 \oplus \ldots \oplus Kv_n$ de rang $n - 1$. Le module quotient $\overline{V} = V/V'$ est un module libre cyclique à générateur $\overline{v_1} = v_1 + V'$. Il contient le sous-module $\overline{U} = (U + V')/V'$. Si $\overline{U} = 0$, alors $U \subset V'$, et l'assertion du théorème est vraie par hypothèse de récurrence. Si au contraire $\overline{U} \neq 0$, le raisonnement développé plus haut pour le cas où $n = 1$ montre que $\overline{U}$ est cyclique à générateur $\overline{u_1} = u_1 + V'$, où $u_1 \in U$. Si de plus $U \cap V' = 0$, alors $u \in U \Rightarrow \overline{u} = u + V' \in \overline{U} \Rightarrow \overline{u} = a_1\overline{u}_1$, $a_1 \in K \Rightarrow u - a_1u_1 \in V' \Rightarrow u = a_1u_1 \Rightarrow U = Ku_1$ est un module libre de rang 1.

Soit enfin $U \cap V' \neq 0$. Par récurrence, le sous-module $U \cap V'$ du module libre V' de rang $n - 1$ possède une base libre $\{u_2, \ldots, u_m\}$, où $0 < m - 1 \leqslant n - 1$. En reprenant presque mot pour mot le raisonnement développé plus haut, nous nous assurons que $\{u_1, u_2, \ldots, u_m\}$ est une K-base libre pour U. En effet, $u \in U \Rightarrow \overline{u} = u + V' \in \overline{U} \Rightarrow \overline{u} = a_1\overline{u}_1$, $a_1 \in K \Rightarrow u - a_1u_1 \in U \cap V' \Rightarrow u - a_1u_1 = a_2u_2 + \ldots + a_mu_m \Rightarrow u = a_1u_1 + a_2u_2 + \ldots + a_mu_m$, $m \leqslant n$. Suivant la proposition 3 (ii) il faut montrer que les générateurs $u_1, \ldots, u_m$ sont linéairement indépendants. Or, $\sum x_iu_i = 0 \Rightarrow x_1\overline{u}_1 = -\sum_{i>0} x_i\overline{u}_i = 0$ dans $\overline{V}$. Par conséquent, $x_1 = 0$, car $\overline{u}_1$ est une base de $\overline{U}$, et puisque $\{u_2, \ldots, u_m\}$ est une base libre de $U \cap V'$, on a $x_2u_2 + \ldots + x_mu_m = 0 \Rightarrow x_2 = \ldots = x_m = 0$. ∎

Corollaire. — *Tout sous-module d'un module de type fini sur un anneau principal est lui-même un module de type fini.*

Démonstration. — Elle résulte des théorèmes 1, 2 et du deuxième théorème d'isomorphie (théorème sur la correspondance entre les sous-modules). ∎

Il n'est pas trop difficile d'obtenir une description complète des modules de type fini sur un anneau principal K. Pourtant, les facteurs décisifs qui stimulent généralement une telle description (modules périodiques sur $\mathbb{Z}$ et sur $P[X]$, voir exemples 1 et 3) ont cessé d'agir (voir chap. 7, § 5 et Annexe) ; quant à l'approche modulaire unique de divers problèmes, le lecteur pourra consulter les ouvrages indiqués dans la bibliographie.

3. Eléments entiers d'un anneau. — Soit K un anneau intègre. On dit qu'un élément $t \in K$ est *entier* (*entier sur* $\mathbb{Z}$) si t est racine du polynôme unitaire $X^n + a_1X^{n-1} + \ldots + a_n \in \mathbb{Z}[X]$. Dans le cas où K est une extension algébrique finie du corps $\mathbb{Q}$, ou K est un corps engendré par tous les nombres algébriques complexes, on parle des *entiers algébriques* en y rangeant naturellement tous les éléments appartenant à $\mathbb{Z}$. L'exercice 9 du chapitre 6, § 4 montre qu'*un*

nombre rationnel t est un entier algébrique si, et seulement si, $t \in \mathbb{Z}$ Si ensuite $a_0u^n + a_1u^{n-1} + \ldots + a_n = 0$, $a_i \in \mathbb{Z}$, alors $(a_0u)^n + a_0a_1(a_0u)^{n-1} + \ldots + a_0^n a_n = 0$, ce qui signifie que *tout nombre algébrique multiplié par un élément approprié* $a_0 \in \mathbb{Z}$ *devient un entier algébrique.*

En revenant au cas général, remarquons qu'il est commode d'interpréter K comme un $\mathbb{Z}$-module. Quels que soient les éléments $t_1, t_2, \ldots, t_n \in K$, ils engendrent dans K un sous-module $Zt_1 + Zt_2 + \ldots + Zt_n$ de type fini. En particulier, si t est un élément entier et $t^n + a_1t^{n-1} + \ldots + a_n = 0$, $a_i \in \mathbb{Z}$, le sous-anneau $\mathbb{Z}[t] \subset K$ est un $\mathbb{Z}$-module de type fini car $\mathbb{Z}[t] = \mathbb{Z}1 + \mathbb{Z}t + \ldots + \mathbb{Z}t^{n-1}$. Réciproquement, soit $\mathbb{Z}[t]$ un $\mathbb{Z}$-module de type fini ayant pour générateurs $v_1, \ldots, v_n \in K$. Alors les relations

$$tv_i = a_{i1}v_1 + a_{i2}v_2 + \ldots + a_{in}v_n, \quad 1 \leqslant i \leqslant n,$$

avec la matrice $A = (a_{ij}) \in M_n(\mathbb{Z})$ permettent de conclure que le système linéaire homogène

$$\begin{array}{l}(t - a_{11})x_1 - a_{12}x_2 - \ldots - a_{1n}x_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ -a_{n1}x_1 - a_{n2}x_2 - \ldots + (t - a_{nn})x_n = 0\end{array}$$

considéré sur le corps des quotients $Q(K)$ possède une solution non nulle $(x_1, \ldots, x_n) = (v_1, \ldots, v_n)$ (tous les v_i ne sont pas simultanément nuls parce que $1 \in \mathbb{Z}[t]$). Par conséquent, le déterminant du système est nul (voir chap. 3) et t est racine du polynôme unitaire $f(T) = \det(TE - A)$. Nous avons ainsi démontré qu'*un élément* $t \in K$ *est entier si, et seulement si, le sous-anneau* $\mathbb{Z}[t] \subset K$ *est un* $\mathbb{Z}$*-module de type fini.*

THÉORÈME 3.—*Les éléments entiers d'un anneau K forment un sous-anneau de K.*

DEMONSTRATION.— Soient $u, v \in K$ des éléments entiers. Alors $\mathbb{Z}[u, v] = \sum_{i \leqslant n,\, j \leqslant m} \mathbb{Z}u^iv^j$ est un $\mathbb{Z}$-module de type fini. Puisque $\mathbb{Z}$ est un anneau principal, le corollaire du théorème 2 (ainsi qu'une vérification directe) montre que les sous-modules $\mathbb{Z}[u - v]$, $\mathbb{Z}[uv]$ sont, eux aussi, des $\mathbb{Z}$-modules de type fini. Suivant le critère indiqué plus haut, les éléments $u - v$, uv doivent être entiers. ■

EXEMPLE.—La racine n-ième ε de l'unité, avec n entier naturel quelconque, est évidemment un entier algébrique. D'après le théorème 3, les combinaisons linéaires à coefficients entiers des racines de l'unité seront aussi des entiers algébriques. En particulier (voir démonstration de la proposition du chap. 8, § 4), *les valeurs* $\chi_\Phi(g)$, $g \in G$, *du caractère* χ_Φ *de toute représentation linéaire* Φ *sur* $\mathbb{C}$ *d'un groupe fini G sont des entiers algébriques.*

4. Suites unimodulaires de polynômes.— Soit $K = P[X_1, \ldots, X_n]$ l'anneau des polynômes à n indéterminées sur un corps commutatif P. On dit que la suite $[f_1, \ldots, f_r]$ de r polynômes $f_i \in K$ est unimodulaire si $Kf_1 + Kf_2 + \ldots + Kf_r = K$, c'est-à-dire si

$$u_1f_1 + u_2f_2 + \ldots + u_rf_r = 1 \qquad (1)$$

pour certains $u_i \in K$, $1 \leqslant i \leqslant r$. Soit ensuite V un module de type fini sur K. En analysant certains problèmes délicats de la géométrie algébrique, le mathématicien français J.-P. Serre a formulé (en 1955) l'hypothèse suivante :

$$\{V \oplus K^s \cong K^{s+t}\} \Rightarrow V \cong K^t$$

qui a été mise sous la forme bien élégante : « toute relation (1) peut s'écrire sous la forme d'une égalité

$$\begin{vmatrix} f_1 & f_2 & \ldots & f_r \\ u_{21} & u_{22} & \ldots & u_{2r} \\ \cdot & \cdot & \cdot & \cdot \\ u_{r1} & u_{r2} & \ldots & u_{rr} \end{vmatrix} = 1 \qquad (2)$$

pour des $u_{ij} \in K$ approchés».

Malgré sa simplicité apparente, cette assertion n'a été démontrée qu'en 1976 par A. A. Souslin (U.R.S.S.) et indépendamment de lui par D. Quillen (Etats-Unis), bien que le cas où $n = 1$ soit attribué encore à Hermite (1848). On a aussi le théorème suivant :

THÉORÈME 4.— *Soient* $a_1, a_2, \ldots, a_r$ $(r \geqslant 2)$ *des éléments non nuls d'un anneau principal* K *et* $d =$ P.G.C.D. $(a_1, \ldots, a_r)$. *Alors il existe une matrice* $A \in M_r(K)$ *ayant pour sa première ligne* $(a_1, a_2, \ldots, a_r)$ *et de* $\det A = d$.

DÉMONSTRATION.— Nous nous appuyons sur les résultats énoncés à la fin du nº 1 du § 2. Pour $r = 2$, en écrivant d sous la forme $d = u_1a_1 + u_2a_2$, avec $u_i \in K$, nous trouvons tout de suite la matrice requise

$$A = \begin{Vmatrix} a_1 & a_2 \\ -u_2 & u_1 \end{Vmatrix}.$$

En raisonnant maintenant par récurrence sur r, représentons $d' =$ P.G.C.D. $(a_1, \ldots, a_{r-1})$ sous la forme $d' = \det A'$, où $A' \in M_{r-1}(K)$ est une matrice ayant pour la première ligne $(a_1, \ldots, a_{r-1})$. Puisque $d =$ P.G.C.D. (d', a_r), on a $d = ud' + va_r$.

Introduisons la matrice

$$A=\left\|\begin{array}{ccc:c} & & & a_r \\ & A' & & 0 \\ & & & \vdots \\ & & & 0 \\ \hdashline \frac{-va_1}{d'}, & \ldots, & \frac{-va_{r-1}}{d'} & u \end{array}\right\|.$$

La première ligne de la matrice $A \in M_n(K)$ est $(a_1, a_2, \ldots, a_r)$. En développant det A suivant les éléments de la dernière colonne nous obtenons

$$\det A = u \det A' + (-1)^{r+1} a_r \det A'' = ud' + a_r (-1)^{r+1} \det A'', \quad (3)$$

où A'' est la matrice extraite de A en supprimant la première ligne et la dernière colonne.

D'autre part, si l'on multiplie la première ligne de la matrice A' par $-v$ et si l'on met le résultat à la dernière place dans A', en gardant l'ordre de toutes les autres lignes (on y arrive par des permutations successives), on obtiendra une matrice A''' qui est la même que dans le cas où la dernière ligne de A'' est multipliée par d'. Il s'ensuit que

$$d' \det A'' = \det A''' = (-1)^{r-1} v \det A' = (-1)^{r-1} vd'.$$

Divisons cette expression par d' et portons le résultat dans (3). Nous obtenons la représentation requise:

$$\det A = ud' + a_r (-1)^{r+1} \det A'' = ud' +$$
$$+ a_r (-1)^{r+1} (-1)^{r-1} v = ud' + va_r = d. \blacksquare$$

Dans le cas de $n > 1$, l'idée maîtresse consiste à examiner l'opération du groupe GL $(r, P[X_1, \ldots, X_{n-1}])$ sur l'ensemble des suites unimodulaires et à raisonner par récurrence sur n. Pour la démonstration on peut consulter l'article original (Souslin A. A. « Проективные модули над кольцом многочленов свободны», (Les modules projectifs sur un anneau des polynômes sont libres), ДАН СССР, **229**, n° 5 (1976), 1063-1066) ou le rapport présenté au Séminaire de N. Bourbaki (Ferrand D., Sém. N. Bourbaki, 28-ième année, 1975/76, Juin, 1976). Son exposé est élémentaire. Pour pouvoir juger du prix qu'on a dû mettre pour y arriver, on peut se reporter à un rapport plus ancien fait également au Séminaire de N. Bourbaki (Bass. H., Sém. N. Bourbaki, 26-ième année, 1973/74, Juin, 1974). Dans les ouvrages cités ci-dessus on trouve les énoncés des problèmes non résolus. Toutes ces questions conviennent très bien à être discutées au cours d'un séminaire spécial.

§ 4. Algèbres sur un corps commutatif

1. Définitions et exemples d'algèbres.— Jusqu'ici nous n'avons pas porté d'attention particulière au fait que la presque totalité des anneaux que nous connaissons, sont en même temps munis d'une structure d'espace vectoriel sur un corps commutatif.

DÉFINITION. — *On appelle algèbre* (*ou algèbre linéaire*) *sur un corps commutatif* P *un couple composé d'un anneau* $(A, +, \cdot)$ *et d'un espace vectoriel* A *sur* P (*l'ensemble de base* A *est le même pour l'anneau et pour l'espace vectoriel*; *l'opération d'addition* $+$ *et l'élément zéro* 0 *sont aussi les mêmes*). *Ceci étant, on a*

$$\lambda (xy) = (\lambda x) y = x (\lambda y)$$

quels que soient $\lambda \in P$, $x, y \in A$. *L'algèbre est dite associative si l'anneau* $(A, +, \cdot)$ *est associatif. La dimension sur* P *de l'espace vectoriel* A *s'appelle aussi dimension de l'algèbre* A.

Les notions fondamentales de la Théorie des anneaux sont aussi applicables, avec les précisions peu signifiantes, aux algèbres. C'est ainsi que par *sous-algèbre* d'une algèbre A on entend tout sous-anneau $B \subset A$ qui est en même temps un sous-espace de l'espace vectoriel A. Si T est un sous-ensemble de A, la sous-algèbre $P[T]$ qu'il engendre est intersection de toutes les sous-algèbres de A contenant T. On définit de manière analogue les idéaux et les *algèbres quotients*. Les *homomorphismes des algèbres* sont des homomorphismes des anneaux, qui sont en même temps des applications P-linéaires.

Le *centre* $Z(A)$ *d'une algèbre associative* A est défini comme ensemble de tous les éléments $a \in A$ commutables à tout élément de A : $a \in Z(A) \Leftrightarrow ax = xa$, $\forall x \in A$. Il est évident que $(a - a') x = ax - a'x = xa - xa' = x (a - a')$, $(aa') x = a (a'x) = a (xa') = (ax) a' = x (aa')$, $(\lambda a) x = \lambda (ax) = \lambda (xa) = x (\lambda a)$, quels que soient $a, a' \in Z(A)$, $\lambda \in P$. Par conséquent, le centre $Z(A)$ est une sous-algèbre de A. L'égalité $Z(A) = A$ a lieu si, et seulement si, A est une algèbre commutative.

Si A est une algèbre associative ayant un élément unité 1, on vérifie directement que $\lambda \cdot 1 \in Z(A)$ et que la relation $\lambda \mapsto \lambda \cdot 1$, $\forall \lambda \in P$, définit une application monomorphe de P dans A. En ce sens, par algèbre A on peut entendre l'anneau A avec le sous-corps prélevé contenu dans le centre $Z(A)$.

Indiquons quelques exemples d'algèbres.

1) L'extension $F \supset P$ de degré fini $[F: P]$ d'un corps commutatif P est manifestement une algèbre associative et commutative (avec élément unité) de dimension finie $\dim_P F = [F : P]$. Nous avons déjà utilisé ce fait au § 1.

2) L'anneau des polynômes $K = P[X_1, \ldots, X_n]$ à coefficients dans un corps commutatif P est muni d'une structure naturelle

d'algèbre associative et commutative de dimension infinie sur le corps P. Remarquons que

$$K = K_0 \oplus K_1 \oplus K_2 \oplus \ldots$$

est une somme directe des sous-espaces vectoriels de dimension finie K_m des polynômes homogènes de degré m ($K_0 = P$) et que $K_i K_j \subset \subset K_{i+j}$. Les algèbres de ce type sont dites *graduées*.

3) L'algèbre commutative $X_{\mathbb{C}}(G)$ ayant un élément unité χ_1, engendrée sur $\mathbb{C}$ par tous les caractères d'un groupe fini G, est de dimension r égale au nombre de classes des éléments conjugués dans G (chap. 8, § 7, théorème 2).

4) L'anneau $M_n(P)$ des matrices carrées d'ordre n à coefficients dans un corps commutatif P est une algèbre de dimension n^2 sur P. Les éléments de base $\{E_{ij} \mid i, j = 1, 2, \ldots, n\}$ de l'algèbre $M_n(P)$ sont multipliés suivant la règle $E_{ik}E_{lj} = \delta_{kl}E_{ij}$. En vertu du théorème 3 du chapitre 2, § 3, le centre $Z(M_n(P)) = \{\lambda E\} \cong P$.

Nous dirons qu'une algèbre associative A avec élément unité est *simple centrale sur un corps commutatif* P si $Z(A) \cong P$ et si A ne contient pas d'idéaux bilatères qui soient différents de 0 et de A.

PROPOSITION 1. *$M_n(P)$ est une algèbre simple centrale.*

Soit J un idéal de $M_n(P)$ distinct de l'idéal zéro et soit

$$0 \neq a = \sum \alpha_{ij}E_{ij} \in J.$$

Si $\alpha_{kl} \neq 0$, alors $E_{st} = \alpha_{kl}^{-1}E_{sk} \cdot a \cdot E_{lt} \in J$, quels que soient $s, t = 1, \ldots, n$, et donc $J = M_n(P)$. ■

Une assertion analogue est vraie pour l'algèbre complète des matrices $M_n(D)$ sur un corps arbitraire D. Le *théorème* extrêmement important *de Wedderburn* (et, dans un contexte plus général, le *théorème de Wedderburn-Artin*) dit que, réciproquement, *toute algèbre associative simple de dimension finie sur un corps commutatif P est isomorphe à $M_n(D)$, où n est un entier naturel défini de façon unique, alors que le corps D (qui est une algèbre de dimension finie sur P) est défini à un isomorphisme près.*

L'algèbre de matrices $M_n(P)$ vérifie encore une propriété universelle suivante :

PROPOSITION 2.— *Toute algèbre associative A de dimension n sur un corps commutatif P est isomorphe à une sous-algèbre de $M_k(P)$, où $k \leqslant n + 1$.*

DÉMONSTRATION.— Supposons que A est une algèbre ayant un élément unité 1. Plongeons-la dans $M_n(P)$. A cet effet, à tout élément $a \in A$ faisons correspondre un opérateur linéaire $L_a : x \mapsto ax$ sur l'espace vectoriel A. La linéarité de L_a résulte du fait que l'opération de multiplication dans A est bilinéaire. Les relations évi-

dentes $L_{\lambda a} = \lambda L_a$, $L_{a+b} = L_a + L_b$, $L_{ab} = L_a L_b$ (associativité!) et $L_1 = \mathcal{E}$, impliquent que l'application φ: $a \mapsto L_a$ est un homomorphisme. Son injectivité est assurée par l'existence de l'élément unité: $a \neq 0 \Rightarrow L_a \cdot 1 = a \cdot 1 = a$, $L_a \neq 0$.

Soit maintenant A une algèbre sans élément unité. Introduisons l'espace vectoriel $\widetilde{A} = P \oplus A$ et définissons sur cet espace l'opération de multiplication en posant $(\lambda, a)(\lambda', a') = (\lambda\lambda', aa' + \lambda a' + \lambda' a)$. On vérifie sans peine que, munie de cette loi de multiplication, $\widetilde{A}$ est une algèbre sur P, ayant un élément unité $(1, 0)$.

Puisque $\dim_P \widetilde{A} = \dim_P A + 1 = n + 1$, le raisonnement précédent permet de plonger $\widetilde{A}$, et donc A, dans $M_{n+1}(P)$. ■

Il n'est pas difficile de constater que la démonstration de la proposition 2 présente une analogie complète avec celle du théorème de Cayley pour les groupes finis. Dans les deux cas on utilise la représentation régulière. Plus généralement, par *représentation d'une algèbre* A sur P on entend tout homomorphisme $A \to \mathfrak{L}(V) = \mathrm{End}_F(V)$, où $F \supset P$ est une extension du corps commutatif P. En d'autres termes, l'espace vectoriel V sur F est muni d'une structure de A-module à gauche, au sens des définitions données au § 3, et l'on a en outre

$$(\lambda x) \cdot v = x \cdot (\lambda v), \quad \forall \lambda \in P, \quad x \in A, \quad v \in V.$$

En choisissant dans V une base quelconque, nous obtiendrons, de même que dans le cas des groupes, une représentation matricielle $A \to M_r(F)$, où $r = \dim_F(V)$.

2. Algèbres à division (corps).— Comme le montre le théorème de Wedderburn énoncé plus haut, l'étude des algèbres à division est une branche importante de la théorie générale des structures des algèbres associatives. Le lemme de Schur (proposition 3 du § 3) confirme, lui aussi, cette considération. Avant d'indiquer des résultats quelconques relatifs aux algèbres à division, analysons une proposition auxiliaire:

PROPOSITION 3.— *Dans une algèbre associative A (ayant un élément unité 1) de dimension n sur un corps commutatif P, tout élément $a \in A$ est racine d'un polynôme (minimal) $f_a \in P[X]$ de degré $\leqslant n$. L'élément $a \in A$ est inversible si, et seulement si, $f_a(0) \neq 0$. Si A ne possède pas de diviseurs de zéro, A est une algèbre à division. Ceci étant, si le corps P est algébriquement clos, alors $n = 1$ et $A = P$.*

DÉMONSTRATION.— Etant donné que A est de dimension finie, les éléments $1, a, a^2, \ldots$ ne peuvent pas être tous linéairement indépendants sur P. Par conséquent, il existe un polynôme unitaire $f_a(X) = X^m + \alpha_1 X^{m-1} + \ldots + \alpha_m \neq 0$ de degré minimal $m \leqslant n$, à coefficients $\alpha_i \in P$, tel que $f_a(a) = 0$. Si $\alpha_m \neq 0$, la relation $f_a(a) = 0$ mise sous la forme $[-\alpha_m^{-1}(a^{m-1} + \alpha_1 a^{m-2} + \ldots$

$\ldots + \alpha_{m-1})] \, a = 1$, montre que a est un élément inversible. Réciproquement, supposons que $a \in A$ n'est pas un diviseur de zéro, mais $\alpha_m = 0$. Alors

$$(a^{m-1} + \alpha_1 a^{m-2} + \ldots + \alpha_{m-1}) \, a = 0 \Rightarrow$$
$$\Rightarrow a^{m-1} + \alpha_1 a^{m-2} + \ldots + \alpha_{m-1} = 0,$$

ce qui est en contradiction avec le fait que $f_a(X)$ est minimal. Donc, $\alpha_m \neq 0$. En particulier, tous les éléments de A qui ne sont pas diviseurs de zéro, sont inversibles.

Si le corps P est algébriquement clos, alors $f_a(X) = (X - c_1) \ldots \ldots (X - c_m)$, $c_i \in P$, d'où $(a - c_1) \, b = 0$, $b = (a - c_2) \ldots \ldots (a - c_m) \neq 0$. L'absence de diviseurs de zéro dans A ne laisse qu'une seule possibilité : $m = 1$ et $a - c_1 = 0$, $a = c_1 \in P$. Comme ceci est vrai pour tout élément $a \in A$, on a $A = P$. ∎

Nous voyons que les propriétés d'une algèbre à division dépendent fortement du corps de base P. C'est naturel que dans l'histoire des mathématiques ce sont les algèbres à division sur le corps $\mathbb{R}$ des nombres réels qui éveillaient un intérêt particulier des savants. L'existence du corps $\mathbb{C} = \mathbb{R} + i\,\mathbb{R}$ incitait à la recherche d'autres « systèmes hypercomplexes ». Ces recherches ont été couronnées de succès en 1843 lorsque Hamilton a construit sa célèbre algèbre des quaternions réels.

EXEMPLE (algèbre des quaternions $\mathbb{H}$).— Formellement,

$$\mathbb{H} = \mathbb{R} + i\mathbb{R} + j\mathbb{R} + k\mathbb{R},$$

où i, j, k sont des grandeurs multipliées suivant la règle

$$i^2 = j^2 = k^2 = -1, \quad ij = k = -ji, \quad jk = i = -kj, \quad ki = j = -ik.$$

L'élément $x = \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k \in \mathbb{H}$ s'appelle *quaternion*. On vérifie directement que $\mathbb{H}$ est une algèbre associative ayant pour centre $Z(\mathbb{H}) = \mathbb{R}$. Il est judicieux de considérer tout d'abord un modèle de l'algèbre $\mathbb{H}$, à savoir l'ensemble

$$\Phi(\mathbb{H}) = \left\{ \left\| \begin{matrix} a & b \\ -\bar{b} & \bar{a} \end{matrix} \right\| \, \middle| \, a, \, b \in \mathbb{C} \right\} \subset M_2(\mathbb{C}).$$

Un exercice élémentaire relatif aux opérations sur les matrices montre que $\Phi(\mathbb{H})$ est un corps. Un exercice analogue a été analysé au chapitre 5, § 1, lorsque nous avons introduit le corps $\mathbb{C}$. On ne devra pas seulement oublier que la multiplication dans $\Phi(\mathbb{H})$ est non commutative. Suivant les règles de calcul d'une matrice inverse on a

$$\left\| \begin{matrix} a & b \\ -\bar{b} & \bar{a} \end{matrix} \right\|^{-1} = \delta^{-1} \left\| \begin{matrix} \bar{a} & -b \\ \bar{b} & a \end{matrix} \right\|,$$

$$\delta = \det \left\| \begin{matrix} a & b \\ -\bar{b} & \bar{a} \end{matrix} \right\| = a\bar{a} + b\bar{b} \; (\neq 0 \text{ pour } a \neq 0 \text{ ou } b \neq 0).$$

Notons, entre autres, qu'il en résulte que le groupe multiplicatif $\Phi(\mathbb{H})^* = \Phi(\mathbb{H}) \setminus \{0\}$ contient un sous-groupe isomorphe à SU (2) (voir chap. 7, § 1).

En posant

$$q_0=\begin{Vmatrix}1 & 0\\ 0 & 1\end{Vmatrix},\quad q_1=\begin{Vmatrix}i & 0\\ 0 & -i\end{Vmatrix},\quad q_2=\begin{Vmatrix}0 & 1\\ -1 & 0\end{Vmatrix},\quad q_3=\begin{Vmatrix}0 & i\\ i & 0\end{Vmatrix}.$$

nous remarquons que

$$q_s^2 = -q_0,\ s \neq 0;\quad q_1q_2 = q_3 = -q_2q_1,\quad q_2q_3 = q_1 = -q_3q_2,$$
$$q_3 q_1 = q_2 = -q_1q_3.$$

Il est clair que l'application $\Phi : \mathbb{H} \to \Phi(\mathbb{H})$ définie par la correspondance $1 \mapsto q_0$, $\boldsymbol{i} \mapsto q_1$, $\boldsymbol{j} \mapsto q_2$, $\boldsymbol{k} \mapsto q_3$ est une représentation de dimension deux sur $\mathbb{C}$ de l'algèbre des quaternions $\mathbb{H}$. Dans ces conditions, au quaternion x est associée la matrice

$$\Phi_x=\begin{Vmatrix}a & b\\ -\bar{b} & \bar{a}\end{Vmatrix}=\alpha_0q_0+\alpha_1q_1+\alpha_2q_2+\alpha_3q_3,$$

où $a = \alpha_0 + i\alpha_1$, $b = \alpha_2 + i\alpha_3$, $i = \sqrt{-1}$. Les *unités quaternioniques* $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ engendrent dans $\mathbb{H}^*$ le groupe quaternionien Q_8 d'ordre 8, que nous connaissons déjà, alors que la restriction $\Phi \mid_{Q_8}$ donne sa représentation irréductible de dimension deux (voir chap. 7, fin du § 3).

Pour tout quaternion $x = \alpha_0 + \alpha_1\boldsymbol{i} + \alpha_2\boldsymbol{j} + \alpha_3\boldsymbol{k}$ est défini un quaternion conjugué $x^* = \alpha_0 - \alpha_1\boldsymbol{i} - \alpha_2\boldsymbol{j} - \alpha_3\boldsymbol{k}$ (analogue d'un nombre complexe conjugué).

L'opération de conjugaison vérifie les propriétés évidentes:

$$(x + y)^* = x^* + y^*;\quad x^* = x \Leftrightarrow x \in \mathbb{R};\quad x^* = -x \Leftrightarrow \alpha_0 = 0$$

(x est un quaternion « imaginaire pur »). Le produit $xx^* = N(x)$ s'appelle *norme du quaternion x*. En utilisant la correspondance Φ, on constate tout de suite que $(xy)^* = y^*x^*$, $N(xy) = N(x)\,N(y)$, et que $N(x) = \det \Phi(x) = \alpha_0^2 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2$.

La place occupée par les quaternions est mise en évidence par le *théorème* suivant dû à *Frobenius. Il n'existe sur le corps* $\mathbb{R}$ *que trois algèbres associatives à division de dimension finie*: $\mathbb{R}$, $\mathbb{C}$ *et* $\mathbb{H}$. Pour la démonstration (que nous ne donnons pas ici) il est essentiel que le polynôme minimal $f_t(X)$ de tout élément $0 \neq t \notin \mathbb{R}$ de l'algèbre à division D sur $\mathbb{R}$ soit quadratique (voir proposition 3 et théorème 1 du chap. 6, § 4).

En partant des considérations topologiques bien profondes, on a démontré il n'y a pas très longtemps que *toute algèbre à division (non nécessairement associative) de dimension finie sur* $\mathbb{R}$ *est de dimension* 1, 2, 4 *ou* 8. Toutes ces possibilités sont réalisables.

Il y a plus de 70 ans, Wedderburn a obtenu un joli résultat relatif aux corps finis, qui a une grande importance pour la géométrie. Nous allons maintenant démontrer ce théorème qui est en rapport direct avec le contenu du § 1.

THÉORÈME 1 (de Wedderburn).— *Tout corps fini est commutatif.*

DÉMONSTRATION.— Soient D un corps fini et Z son centre. Il est évident que Z est un corps commutatif, et D est un espace vectoriel

de dimension finie sur Z:

$$D = Ze_1 \dot{+} Ze_2 \dot{+} \dots \dot{+} Ze_n.$$

Conformément aux résultats du § 1, $Z = \mathbb{F}_q$ pour un $q = p^m$, si bien que $|D| = q^n$. Soit ensuite $x \in D \setminus Z$. Les éléments qui commutent à x forment un ensemble $C(x) = \{y \in D \mid yx = xy\}$ stable pour les opérations d'addition et de multiplication. En d'autres termes, $C(x)$ est une sous-algèbre à division dans D, contenant Z. Si q^d est le nombre d'éléments appartenant à $C(x)$, alors $d = d(x)$ est un diviseur de n, $d < n$, car en interprétant D comme un espace vectoriel à gauche

$$D = C(x) f_1 \dot{+} \dots \dot{+} C(x) f_r$$

sur $C(x)$, nous avons $q^n = |C(x)|^r = q^{dr}$. Remarquons maintenant que Z^* est le centre du groupe multiplicatif D^* et que $(q^n - 1)/(q^d - 1) = (D^* : C(x)^*)$ est le nombre d'éléments conjugués de x dans D^*. Par conséquent, la formule (2′) du chapitre 7, § 2, peut se mettre sous la forme

$$q^n - 1 = |D^*| = (q-1) + \sum_d \frac{q^n-1}{q^d-1}, \qquad (*)$$

où d parcourt un ensemble des diviseurs de n inférieurs à n. Les propriétés du polynôme cyclotomique $\Phi_n(X)$ établies au § 1 montrent (voir § 1, exercice 6) que le nombre entier $\Phi_n(q)$ divise aussi bien $q^n - 1$ que $(q^n - 1)/(q^d - 1)$, avec $d \mid n$, $d < n$. Dans ces conditions, suivant (*) on a $\Phi_n(q) \mid (q - 1)$, ce qui entraîne (voir § 1, exercice 7) l'égalité $n = 1$ et donc la commutativité de $D = Z$. ∎

3. Algèbres de groupes et modules sur ces algèbres.— En considérant au chapitre 8, § 1 la représentation régulière d'un groupe fini G, nous avons introduit l'espace vectoriel $\langle e_g \mid g \in G\rangle_K$ sur un corps commutatif K. Nous allons maintenant transformer cet espace vectoriel en une K-algèbre, en posant $e_g e_h = e_{gh}$ et en étendant cette loi suivant la linéarité à des « vecteurs » arbitraires $\sum \alpha_g e_g$, $\alpha_g \in K$. Pour simplifier l'écriture, on remplace généralement e_g par g et on considère l'ensemble $K[G]$ de toutes les sommes formelles possibles $\sum \alpha_g g$, $\alpha_g \in K$. Par définition, $\sum \alpha_g g = \sum \beta_g g \Leftrightarrow \alpha_g = \beta_g$, $\forall g \in G$. Les opérations sur les sommes formelles:

$$\begin{gathered} \sum_g \alpha_g g + \sum_g \beta_g g = \sum_g (\alpha_g + \beta_g) g, \\ \lambda \left(\sum \alpha_g g\right) = \sum_g \lambda \alpha_g g, \\ \left(\sum_g \alpha_g g\right)\left(\sum_h \beta_h h\right) = \sum_{g,h} \alpha_g \beta_h gh = \sum \gamma_u u, \\ \text{où } \gamma_u = \sum_g \alpha_g \beta_{g^{-1}u} \end{gathered} \qquad (1)$$

définissent sur $K\,[G]$ une structure d'algèbre associative. On convient de donner à $K\,[G]$ le nom d'*algèbre de groupes du groupe fini* G sur le corps commutatif K. Les éléments de base de l'espace $K\,[G]$ sont les sommes formelles $1\cdot g$, $g \in G$, identifiées aux éléments $g \in G$; $\dim_K K\,[G] = |\,G\,|$. Ainsi, le groupe G est considéré comme étant plongé dans l'algèbre $K\,[G]$. L'élément unité $e \in G$ est en même temps l'élément unité de $K\,[G]$. Dans le cas où K est un anneau associatif et commutatif unitaire, on obtient l'*anneau* $K\,[G]$ du groupe G sur K.

En outre, une construction analogue est applicable à un groupe G arbitraire, non nécessairement fini, si l'on convient de ne considérer que des sommes $\sum \alpha_g g$ ayant un nombre fini de coefficients différents de zéro. Il est aussi commode d'interpréter $A = \sum \alpha_g g$ comme une fonction sur le groupe G (à valeurs $A\,(g) = \alpha_g$ dans K) égale presque partout à zéro (c'est-à-dire admettant un nombre fini de valeurs non nulles). Ceci étant, aux formules (1) correspondent l'opération d'addition :

$$(A_1 + A_2)\,(g) = A_1\,(g) + A_2\,(g)$$

et la *convolution des fonctions*:

$$A_3 = A_1 * A_2, \qquad A_3\,(u) = \sum_g A_1\,(g)\,A_2\,(g^{-1}u).$$

La théorie des anneaux de groupes est une branche bien importante de l'algèbre, qui étudie ses problèmes spécifiques, mais pour nous, $K\,[G]$ n'est qu'une illustration des notions générales introduites dans les deux derniers chapitres.

THÉORÈME 2. — *Il existe une correspondance biunivoque entre les $K\,[G]$-modules qui sont des espaces vectoriels de dimension finie sur un corps commutatif K, et les représentations linéaires du groupe G.*

DÉMONSTRATION. — Soit (Φ, V) une représentation du groupe G. Prolongeons Φ par linéarité aux éléments de $K\,[G]$, en définissant

$$\tilde{\Phi}\left(\sum \alpha_g g\right) = \sum \alpha_g \Phi\,(g),$$

et posons

$$\left(\sum \alpha_g g\right) \circ v = \sum \alpha_g \Phi\,(g)\,v, \qquad \forall v \in V.$$

L'opération $\circ$ introduit sur V une structure de $K\,[G]$-module au sens habituel de ce terme. Remarquons que

$$\left(\sum \alpha_g g\right) \circ (\lambda v) = \sum \alpha_g \Phi\,(g)\,(\lambda v) = \sum \alpha_g \lambda \Phi\,(g)\,v =$$
$$= \lambda \left(\sum \alpha_g \Phi\,(g)\,v\right) = \lambda \left(\left(\sum \alpha_g g\right) \circ v\right),$$

c'est-à-dire que les multiplications par des scalaires dans V et dans $K\,[G]$ sont compatibles. Le couple $(\tilde{\Phi}, V)$ peut s'appeler naturellement représentation linéaire de l'algèbre $K\,[G]$.

Réciproquement, si V est un espace vectoriel sur K, qui est un module sur $K[G]$ avec l'opération $(\sum \alpha_g g, v) \mapsto (\sum \alpha_g g) \circ v$, alors, en posant

$$\tilde{\Phi}(\sum \alpha_g g)\, v = (\sum \alpha_g g) \circ v,$$

nous définirons un homomorphisme $\tilde{\Phi}: K[G] \to \mathrm{End}_K(V)$ (c'est-à-dire une représentation de l'algèbre $K[G]$) dont la restriction $\Phi = \tilde{\Phi}|_G$ à G donne une représentation du groupe G. ∎

Eu égard au théorème 1, l'espace de représentation V du groupe G ou tout simplement *G-module*. Des modifications terminologiques correspondantes concernent aussi d'autres notions adoptées en Théorie des représentations.

Soient G un groupe fini et $K = \mathbb{C}$ le corps des nombres complexes. D'après les résultats du chapitre 8, tout G-module irréductible sur $\mathbb{C}$ (c'est-à-dire un $\mathbb{C}[G]$—module) de caractère χ_i est isomorphe à un idéal à gauche J_i de l'algèbre $\mathbb{C}[G]$ (voir à ce propos § 3, exemple 4). Si $\dim_{\mathbb{C}} J_i = n_i$, alors $\mathbb{C}[G]$ contient une somme directe $A_i = J_{i,1} \oplus \ldots \oplus J_{i,n_i}$ de n_i idéaux à gauche $\mathbb{C}[G]$-isomorphes à $J_i = J_{i,1}$. En choisissant dans chaque classe d'idéaux à gauche isomorphes un représentant J_i, nous pouvons écrire une décomposition

$$\mathbb{C}[G] = A_1 \oplus A_2 \oplus \ldots \oplus A_r, \qquad (2)$$

qui correspond à la décomposition de la représentation régulière du groupe G. Remarquons que chacune des composantes A_i est définie de façon unique.

Maintenant, si J est un idéal à gauche minimal de l'algèbre $\mathbb{C}[G]$, et $t \in \mathbb{C}[G]$, alors Jt est aussi un idéal à gauche minimal (ou l'idéal nul). Par suite, l'application $\varphi: J \to Jt$ définie par la correspondance $v \mapsto vt$ $(v \in J)$ est soit l'application nulle, soit un $\mathbb{C}[G]$-isomorphisme, car $xv \in J$ pour tout $x \in \mathbb{C}[G]$ et $\varphi(xv) = (xv)\,t = x(vt) = x\varphi(v)$. Pour cette raison, $J \subset A_i \Rightarrow Jt \subset A_i$, $\forall t \in \mathbb{C}[G]$, et donc A_i est un idéal bilatère de $\mathbb{C}[G]$. La décomposition (2) est directe, de sorte que

$$i \neq j \Rightarrow A_i A_j \subset A_i \cap A_j = 0.$$

Nous allons essayer d obtenir une information plus précise sur la décomposition (2) en nous appuyant sur la théorie des caractères développée au chapitre 8. Cherchons d'abord le centre $Z(\mathbb{C}[G])$ de l'algèbre de groupes $\mathbb{C}[G]$. Per définition

$$z \in Z(\mathbb{C}[G]) \Leftrightarrow zg = gz, \quad \forall g \in G.$$

Si $z = \sum_{h \in G} \gamma_h h$, alors

$$\sum_{t \in G} \gamma_{g^{-1}t}\, t = g\,(\sum_h \gamma_h h) = (\sum_h \gamma_h h)\, g = \sum_{t \in G} \gamma_{tg^{-1}}\, t,$$

d'où $\gamma_{g^{-1}t} = \gamma_{tg^{-1}}$, $\forall t \in G$. En posant $t = gh$, on obtient $\gamma_h = \gamma_{ghg^{-1}}$. Cela signifie que

$$Z(\mathbb{C}[G]) = \langle z_1, z_2, \ldots, z_r \rangle_{\mathbb{C}},$$

où

$$z_i = \sum_{g \in g_i^G} g; \quad i = 1, 2, \ldots, r \tag{3}$$

($g_1, g_2, \ldots, g_r$ sont les représentants des classes des éléments conjugués du groupe G). Il est clair que $z_1, z_2, \ldots, z_r$ sont des éléments linéairement indépendants et donc $\dim_{\mathbb{C}} Z(\mathbb{C}[G]) = r$.

A tout élément $a \in A_i$ faisons correspondre un opérateur linéaire $L_a^{(i)}$ agissant sur l'idéal à gauche minimal $J_i = J_{i,1}$ suivant la règle $L_a^{(i)}(v) = av$, $v \in J_i$. Comme il est évident que $L_{\lambda a}^{(i)} = \lambda L_a^{(i)}$, $L_{a+b}^{(i)} = L_a^{(i)} + L_b^{(i)}$, $L_{ab}^{(i)} = L_a^{(i)} L_b^{(i)}$, $\varphi: \ a \to L_a^{(i)}$ est un homomorphisme de l'algèbre A_i dans l'algèbre des endomorphismes $\mathrm{End}_{\mathbb{C}} J_i \cong \cong M_{n_i}(\mathbb{C})$. Supposons que $0 \neq a \in \mathrm{Ker}\, \varphi$, c'est-à-dire $aJ_i = 0$. Tous les idéaux à gauche $J_{i,j}$ sont $\mathbb{C}[G]$-isomorphes, et si $\varphi_j: J_i \to \to J_{i,j}$ est un isomorphisme, alors

$$aJ_{i,j} = a\varphi_j(J_i) = \varphi_j(aJ_i) = \varphi_j(0) = 0.$$

Par suite, $aA_i = aJ_{i,1} + \ldots + aJ_{i,n_i} = 0$, et dans ce cas on a aussi $a\,\mathbb{C}[G] = 0$ car $a \in A_i \Rightarrow aA_j = 0$ pour tout $j \neq i$. Pourtant $ae = a \neq 0$. Cette contradiction prouve que $\mathrm{Ker}\, \varphi = 0$. Par conséquent, φ est un monomorphisme, et comme $\dim A_i = n_i^2 = \dim M_{n_i}(\mathbb{C})$ on a $A_i \overset{\varphi}{\cong} M_{n_i}(\mathbb{C})$. Compte tenu de la proposition 1 on peut énoncer le théorème suivant relatif à la structure de l'algèbre de groupes $\mathbb{C}[G]$:

THÉORÈME 3.— *L'algèbre $\mathbb{C}[G]$ d'un groupe fini G sur le corps $\mathbb{C}$ des nombres complexes se décompose en somme directe* (2) *des idéaux bilatères premiers isomorphes aux algèbres matricielles complètes:*

$$\mathbb{C}[G] \cong M_{n_1}(\mathbb{C}) \oplus M_{n_2}(\mathbb{C}) \oplus \ldots \oplus M_n(\mathbb{C}).$$

En particulier, l'algèbre de groupes d'un groupe abélien d'ordre n sur C est isomorphe à la somme directe de r exemplaires du corps $\mathbb{C}$. ■

COROLLAIRE. (théorème de Burnside). —*Soit Φ une représentation matricielle irréductible de degré n sur $\mathbb{C}$ d'un groupe fini G. Alors, parmi les matrices Φ_g, $g \in G$, il existe n^2 matrices linéairement indépendantes, c'est-à-dire $\langle \Phi_g \mid g \in G \rangle_{\mathbb{C}} = M_n(\mathbb{C})$.* ■

La structure du centre $Z(\mathbb{C}[G])$ en tant que sous-algèbre commutative de $\mathbb{C}[G]$ est entièrement définie par les *constantes de struc-*

ture, c'est-à-dire par les entiers n_{ij}^k figurant dans les relations

$$z_i z_j = \sum_{k=1}^{r} n_{ij}^k z_k. \qquad (4)$$

Eu égard à l'expression (3) pour z_i, on se rend compte sans peine que n_{ij}^k est le nombre de couples (g, h), $g \in g_i^G$, $h \in g_j^G$, tels que $gh = g_k$.

Choisissons dans $Z(\mathbb{C}[G])$ une autre base

$$e_i = \frac{n_i}{|G|} \sum_{k=1}^{r} \overline{\chi_i(g_k)}\, z_k = \frac{n_i}{|G|} \sum_{g \in G} \overline{\chi_i(g)}\, g, \quad 1 \leqslant i \leqslant r. \qquad (5)$$

Ici, de même qu'au chapitre 8, § 5, $\chi_1, \ldots, \chi_r$ sont les caractères des représentations irréductibles et $n_1, \ldots, n_r$ leurs degrés. Le passage inverse se fait au moyen de la formule

$$z_k = |g_k^G| \sum_{i=1}^{r} \frac{\chi_i(g_k)}{n_i} e_i.$$

Pour s'en assurer, il convient d'utiliser la relation (4) du chapitre 8, § 5. Elle montre aussi que

$$\sum_{i=1}^{r} e_i = \frac{1}{|G|} \sum_{g \in G} g \sum_i n_i \overline{\chi_i(g)} = \frac{1}{|G|} \sum_{g \in G} g \sum \chi_i(e)\, \overline{\chi_i(g)} =$$
$$= \frac{1}{|G|} e\, |C_G(e)| = e.$$

En appliquant la relation généralisée d'orthogonalité (voir chap. 8, § 4, exercice 1), on trouve

$$e_i e_j = \frac{n_i n_j}{|G|^2} \sum_{g,\, t \in G} \overline{\chi_i(g)}\, \overline{\chi_j(t)}\, gt =$$
$$= \frac{n_i n_j}{|G|^2} \sum_{h \in G} \left\{ \frac{1}{|G|} \sum_{g \in G} \overline{\chi_i(g)}\, \chi_j(hg) \right\} h^{-1} =$$
$$= \frac{n_i n_j}{|G|} \frac{\delta_{ij}}{n_i} \sum \chi_i(h)\, h^{-1} = \delta_{ij} e_i.$$

Ainsi, les éléments de centre e_i calculés à l'aide de la formule (5) satisfont aux relations

$$e = e_1 + e_2 + \ldots + e_r,$$
$$e_i^2 = e_i, \quad e_i e_j = 0, \quad i \neq j, \qquad (6)$$

et s'appellent pour cette raison (et conformément à la tradition) *idempotents orthogonaux centraux* de l'algèbre de groupes $\mathbb{C}[G]$. La

relation $e = e_1 + \ldots + e_r$ exprime la condition de plénitude de ce système. En posant $B_i = e_i\mathbb{C}[G]$, nous constatons tout de suite, que B_i est un idéal bilatère dans $\mathbb{C}[G]$ ayant un élément unité e_i et qu'il existe une décomposition en somme directe

$$\mathbb{C}[G] = B_1 \oplus B_2 \oplus \ldots \oplus B_r. \qquad (7)$$

De (5) il résulte immédiatement que

$$\chi_j(e_i) = n_i \frac{1}{|G|} \sum_g \overline{\chi_i(g)}\, \chi_j(g) = n_i \delta_{ij}.$$

C'est pourquoi B_i contient un idéal à gauche minimal $J \subset A_i$ correspondant au caractère χ_i. Comme A_i et B_i sont des idéaux bilatères, $A_i \subset B_i$. En comparant les décompositions (2) et (7) nous concluons que $A_i = B_i$. Ainsi, nous avons démontré une variante plus perfectionnée du théorème 3.

THÉORÈME 4.— *Les éléments e_i, $1 \leqslant i \leqslant r$, calculés au moyen de la formule* (5), *forment un système plein d'idempotents orthogonaux centraux de l'algèbre* $\mathbb{C}[G]$ *du groupe fini G. La composante simple* $e_i\mathbb{C}[G]$ *de la décomposition directe*

$$\mathbb{C}[G] = e_1\mathbb{C}[G] \oplus e_2\mathbb{C}[G] \oplus \ldots \oplus e_r\mathbb{C}[G],$$

isomorphe à l'algèbre matricielle complète $M_{n_i}(\mathbb{C})$, *contient tous les idéaux à gauche minimaux correspondant au caractère* χ_i. ■

Toute la théorie des représentations des groupes peut être développée à partir du théorème de Weddenburn-Artin (voir n° 1) et de la théorie générale de la structure des algèbres de groupes (sa conclusion finale pour les groupes finis est énoncée au théorème 3). Dans nos raisonnements, nous sommes allés en sens inverse en nous appuyant, au fond, seulement sur le lemme de Schur.

Avant de clore ce numéro, démontrons une assertion utile ayant trait aux degrés de représentations.

THÉORÈME 5.— *Le degré n d'une représentation irréductible* (Φ, V) *sur* $\mathbb{C}$ *d'un groupe fini G divise l'ordre* $|G|$.

DÉMONSTRATION.— Soit $\tilde{\Phi}$ la représentation correspondante de l'algèbre de groupes $\mathbb{C}[G]$. D'après le lemme de Schur (§ 3, proposition 3) l'opérateur linéaire $\tilde{\Phi}(z_i)$ commutable à tous les $\Phi(g)$, $g \in G$, et, de ce fait, appartenant à $\mathrm{End}_{\mathbb{C}[G]}(V)$ doit être multiple de l'opérateur unité : $\tilde{\Phi}(z_i) = \omega_i \mathcal{E}$. On a

$$n\omega_i = \mathrm{tr}\,\omega_i\mathcal{E} = \mathrm{tr}\,\tilde{\Phi}(z_i) = \sum \mathrm{tr}\,\Phi(g_i^h) = |g_i^G|\, \chi_\Phi(g_i),$$

d'où

$$\omega_i = \frac{|g_i^G|\, \chi_\Phi(g_i)}{n}.$$

En appliquant $\widetilde{\Phi}$ aux relations (4), on obtient

$$\omega_i\omega_j = \sum_{k=1}^{r} n_{ij}^k \omega_k.$$

Par suite, $\mathbb{Z}[\omega_i]$ est un sous-module du $\mathbb{Z}$-module $\mathbb{Z}[\omega_1, \ldots, \omega_r]$ de type fini et, conformément aux résultats du § 3, n° 3, ω_i est un entier algébrique. D'après les mêmes résultats

$$\frac{|G|}{n} = \frac{|G|}{n}(\chi_\Phi, \chi_\Phi)_G = \frac{1}{n}\sum \chi_\Phi(g)\overline{\chi_\Phi(g)} =$$

$$= \frac{1}{n}\sum_{i=1}^{r} |g_i^G| \cdot \chi_\Phi(g_i)\overline{\chi_\Phi(g_i)} = \sum \omega_i \overline{\chi_\Phi(g_i)}$$

est un entier algébrique. Donc, $\frac{|G|}{n} \in \mathbb{Z}$. ■

4. Algèbres non associatives.— Soit A une algèbre quelconque (c'est-à-dire non nécessairement associative) de dimension arbitraire sur un corps commutatif P. A tout triplet (x, y, z) d'éléments de A faisons correspondre l'expression $(x, y, z) = (xy)z - x(yz)$ appelée leur *associateur*. Suivant les relations identiques qui lient les associateurs ou d'autres expressions, on obtient divers types d'algèbres (on dit aussi de *classes primitives* ou de *variétés*). Comme exemple on peut indiquer:

1) *les algèbres associatives*: $(x, y, z) = 0$;
2) *les algèbres élastiques*: $(x, y, x) = 0$;
3) *les algèbres alternatives*: $(x, x, y) = (y, x, x) = 0$;
4) *les algèbres de Jordan*: $(x, y, x^2) = 0$; $xy - yx = 0$.

En empruntant cette voie axiomatique on peut, certes, progresser indéfiniment. Pourtant il est remarquable que de nombreuses classes d'algèbres non associatives sont apparues de façon naturelle dans des domaines fort éloignés de l'algèbre en tant que science. Les exemples les plus éclatants sont ceux d'algèbres de Jordan qui sont venues aux mathématiques de la mécanique quantique (du physicien Jordan), et d'algèbres de Lie qui n'étaient destinées primitivement qu'à la description (dans des conditions déterminées) de la structure locale des groupes topologiques (Sophus Lie est mathématicien norvégien du XIXe siècle). Quant aux algèbres de Lie, dont on a fait mention, en passant, dans les pages de ce livre, nous allons les examiner de plus près.

Dans une algèbre de Lie L sur un corps commutatif P, on convient de désigner le produit des éléments $x, y \in L$ par $[xy]$. En vertu de la définition même d'une algèbre de Lie, l'opération bilinéaire $(x, y) \mapsto [xy]$ satisfait à deux conditions:

(i) $[xx] = 0$ ($[xy] = -[yx]$, *anticommutativité*);
(ii) $[[xy]z] + [[yz]x] + [[zx]y] = 0$ (*identité de Jacobi*).

EXEMPLE 1.— Soit A une algèbre associative sur un corps commutatif P. Définissons sur l'espace vectoriel A une structure d'algèbre de Lie $L(A)$ en posant $[xy] = xy - yx$. Il est clair que $[xx] = 0$. On a ensuite

$$[[xy]z] = (xy - yx)z - z(xy - yx) = xyz - yxz - zxy + zyx,$$
$$[[yz]x] = (yz - zy)x - x(yz - zy) = yzx - zyx - xyz + xzy,$$
$$[[zx]y] = (zx - xz)y - y(zx - xz) = zxy - xzy - yzx + yxz.$$

Par une simple addition on obtient l'identité de Jacobi.

En particulier, soit $A = \mathrm{End}_P(V) = \mathfrak{L}(V)$ l'algèbre de tous les opérateurs linéaires d'un espace vectoriel de dimension finie V sur P. Tout homomorphisme $\varphi: L \to L(\mathfrak{L}(V))$ s'appelle *représentation de l'algèbre de Lie L*. L'espace de représentation V s'appelle aussi *L-module* (ou *module sur l'algèbre de Lie L*). Formellement, un L-module est défini par trois axiomes:

(L1) $x(\alpha u + \beta v) = \alpha xu + \beta xv$;

(L2) $(\alpha x + \beta y)v = \alpha xv + \beta yv$;

(L3) $[xy]v = x(yv) - y(xv)$.

EXEMPLE 2.— On appelle *dérivation d'une algèbre* quelconque K (non nécessairement associative) sur un corps commutatif P la *dérivation* $\mathscr{D}$ de l'anneau K (voir définition au chap. 6, § 1, n° 3) commutable à l'opération des constantes de P : $\mathscr{D}(\lambda a) = \lambda\mathscr{D}(a)$, $\lambda \in P$, $a \in K$. L'exemple 1 et l'exercice 8 du chapitre 6, § 1 montrent que la multiplication $[\mathscr{D}_1\mathscr{D}_2] = \mathscr{D}_1\mathscr{D}_2 - \mathscr{D}_2\mathscr{D}_1$ confère à l'ensemble Der (K) qui est un espace vectoriel sur P, une structure d'algèbre de Lie. En particulier si $K = P[X]$ est l'algèbre des polynômes, alors Der (K) se compose de dérivations $\mathscr{D}_u$, $u \in K$, agissant suivant la loi: $\mathscr{D}_u(f) = u\dfrac{df}{dX} = uf'$. Par définition, $[\mathscr{D}_u\mathscr{D}_v](f) = \mathscr{D}_u(\mathscr{D}_v f) - \mathscr{D}_v(\mathscr{D}_u f) = \mathscr{D}_u(vf') - \mathscr{D}_v(uf') = u(vf')' - v(uf')' = u(v'f' + vf'') - v(u'f' + uf'') = (uv' - u'v)f'$. Par conséquent, $[\mathscr{D}_u\mathscr{D}_v] = \mathscr{D}_{uv'-u'v}$, et on voit que l'algèbre Der(K) est isomorphe à l'algèbre de Lie de dimension infinie $(K, [\ \ ])$ ayant K pour espace de base et munie d'opération de multiplication $[uv] = uv' - u'v$. En posant $K_{(i)} = \langle X^{i+1}\rangle_{\mathbb{C}}$, on obtient la décomposition de K en somme directe

$$K = K_{(-1)} \oplus K_{(0)} \oplus K_{(1)} \oplus K_{(2)} \oplus \ldots,$$

qui vérifie la propriété d'une *algèbre de Lie graduée*: $[K_{(i)}K_{(j)}] \subset K_{(i+j)}$ (comparer avec l'exemple 2 du n° 1). L'algèbre de Lie $(K, [\ \ ])$ opère sur l'espace vectoriel K de deux façons: 1) $(a, f) \mapsto af'$ (*opération naturelle*); 2) $(a, f) \mapsto af' - a'f$ (*opération par endomorphismes associés*). Il en résulte deux $(K, [\ \ ])$-modules non isomorphes.

EXEMPLE 3.— Les matrices hermitiennes gauches de trace nulle K_1, K_2, K_3, construites dans l'exercice 3 du chapitre 7, § 1, suivant le groupe SU (2), vérifient les relations

$$[K_1K_2] = K_3,\quad [K_2K_3] = K_1,\quad [K_3K_1] = K_2,$$

qui reproduisent exactement les règles du produit vectoriel des vecteurs dans $\mathbb{R}^3$ ($[K_sK_t] = K_sK_t - K_tK_s$ sont des « commutateurs » des matrices dans $M_2(\mathbb{C})$; voir exemple 1). C'est pourquoi l'espace réel à trois dimensions $\langle K_1, K_2, K_3\rangle_{\mathbb{R}}$ est muni d'une structure d'algèbre de Lie.

De la théorie générale des représentations des groupes compacts il résulte qu'il existe une correspondance biunivoque entre les représentations irréductibles du groupe SU(2) et celles de son algèbre de Lie $\mathbf{su}(2) = \langle K_1, K_2, K_3 \rangle_{\mathbb{R}}$. Intuitivement on peut le réaliser en tenant compte de la continuité de la représentation du groupe et en considérant dans l'enveloppe linéaire des opérateurs $\Phi(g_t)$ (où g_t est un élément du groupe SU(2), qui dépend de $t \in \mathbb{R}$ par dérivation; $g_0 = e$) l'opérateur linéaire $\lim\limits_{t\to 0} \frac{1}{t} \Phi(g_t)$ qui est déjà contenu dans l'algèbre $\mathbf{su}(2)$. Pour prouver que la liste des représentations irréductibles du groupe SU(2) que nous avons obtenues au chapitre 8, § 6, est complète, il nous faut démontrer que pour tout entier naturel n il existe, à un isomorphisme près, un et un seul $\mathbf{su}(2)$-module irréductible de dimension n sur $\mathbb{C}$. A cet effet, il est commode de passer dès le début de l'algèbre de Lie réelle $\mathbf{su}(2)$ à sa « complexification » qui coïncide avec l'algèbre de Lie

$$L = \mathbf{sl}(2) = \mathbf{su}(2) \otimes_{\mathbb{R}} \mathbb{C}$$

de toutes les matrices complexes de dimensions 2×2 de trace nulle. Les éléments de base

$$e_{-1} = -iK_1 + K_2, \quad e_0 = -2iK_3, \quad e_1 = -iK_1 - K_2$$

de l'algèbre L sont multipliés suivant la règle

$$[e_1 e_{-1}] = e_0, \quad [e_0 e_{-1}] = -2e_{-1}, \quad [e_0 e_1] = 2e_1. \tag{8}$$

En oubliant pour un instant l'origine de L, on peut considérer que $L = \langle e_{-1}, e_0, e_1 \rangle_{\mathbb{C}}$ est une algèbre de Lie abstraite de dimension trois sur $\mathbb{C}$ munie de la table de multiplication (8). On vérifie sans peine que L est une algèbre de Lie simple. Par suite, tout L-module irréductible de dimension >1 sera exact.

Soit d'abord $V \neq 0$ un L-module arbitraire de dimension finie sur $\mathbb{C}$ et soient E_{-1}, E_0, E_1 des opérateurs linéaires sur V associés respectivement aux éléments e_{-1}, e_0, e_1. La théorie des représentations des algèbres de Lie a élaboré sa propre terminologie dont nous allons nous en tenir. Le sous-espace propre $V^\lambda = \{v \in V \mid E_0 v = \lambda v\}$ de l'opérateur E_0 dans V, correspondant à la valeur propre $\lambda \in \mathbb{C}$, se compose de vecteurs dont on convient de dire qu'ils sont de *poids* λ. La dimension $\dim V^\lambda$ s'appelle *ordre de multiplicité du poids* λ.

LEMME 1.—*Si $v \in V^\lambda$, alors $E_1 v \in V^{\lambda+2}$, $E_{-1} v \in V^{\lambda-2}$ (E_1 est un opérateur de « majoration » et E_{-1} un opérateur de « minoration »).*

DÉMONSTRATION.— D'après l'axiome (L3) on a

$$E_0(E_1 v) = [E_0 E_1] v + E_1(E_0 v) = 2E_1 v + E_1(\lambda v) = (\lambda + 2) E_1 v,$$

de sorte que, par définition, $E_1 v \in V^{\lambda+2}$. De manière analogue, $E_0(E_{-1}v) = (\lambda - 2) E_{-1}v$. ■

Comme on l'apprend en Algèbre linéaire, les vecteurs correspondant à des valeurs propres différentes sont linéairement indépendants. De ce fait, la somme $W = \sum_\lambda V^\lambda \subset V$ est directe. Du lemme 1, il résulte aussi que $W = \sum_\lambda V^\lambda$ est un L-sous-module de V. Puisque $W \neq 0$, dans le cas d'un L-module irréductible V on doit avoir l'égalité $W = V$.

Un vecteur $v_0 \in V$ sera appelé *vecteur dominant de poids* λ si $v_0 \neq 0$ et $E_1 v_0 = 0$, $E_0 v_0 = \lambda v_0$.

LEMME 2.— *Tout L-module V de dimension finie possède un vecteur dominant.*

DÉMONSTRATION.— Prenons un vecteur arbitraire ($\neq 0$) v de poids μ et construisons une suite de vecteurs v, $E_1 v$, $E_1^2 v$, . . . de poids μ, $\mu + 2$, $\mu + 4$, . . . (voir lemme 1). Puisque $\dim V < \infty$, on a $E_1^{m+1} v = 0$ pour un certain m. En donnant à m sa valeur minimale, on peut poser $v_0 = E_1^m v$, $\lambda = \mu + 2m$. ■

Considérons à titre d'exemple un espace vectoriel V_n de dimension $n + 1$ sur $\mathbb{C}$ ayant une base fixe $v_0, v_1, \ldots, v_n$. Définissons les opérateurs E_{-1}, E_0, E_1 par les formules

$$\begin{aligned} E_{-1} v_m &= (m + 1)\, v_{m+1}, \\ E_0 v_m &= (n - 2m)\, v_m, \\ E_1 v_m &= (n - m + 1)\, v_{m-1}, \end{aligned} \tag{9}$$

en posant $v_{-1} = 0 = v_{n+1}$. Une vérification directe montre que les relations

$$\begin{aligned} E_1 (E_{-1} v_m) - E_{-1} (E_1 v_m) &= E_0 v_m, \\ E_0 (E_{-1} v_m) - E_{-1} (E_0 v_m) &= -2E_{-1} v_m, \\ E_0 (E_1 v_m) - E_1 (E_0 v_m) &= 2E_1 v_m, \end{aligned}$$

qui sont en accord avec la table de multiplication (8) et avec les axiomes du L-module, sont vérifiées. Puisque $E_1 v_0 = (n + 1)\, v_{-1} = 0$, $E_0 v_0 = n v_0$, alors v_0 est un vecteur dominant de poids n, et l'espace V_n tout entier s'écrit sous la forme d'une somme directe

$$V_n = V^n \oplus V^{n-2} \oplus \ldots \oplus V^{-n} \tag{10}$$

des sous-espaces de poids $V^{n-2m} = \langle v_m \rangle$ de dimension un (la multiplicité de chaque poids est 1). En supposant l'existence d'un sous-module $U \neq 0$ de V_n, prenons un vecteur propre quelconque $u \in U$ de l'opérateur E_0. Conformément à la décomposition (10), $u = \lambda v_m$ pour un certain m. L'application successive de l'opérateur de « majoration » E_1 (voir formules (9)) nous donne les relations

$v_{m-1} \in U, \ldots, v_0 \in U$, alors qu'à l'aide de l'opérateur de « minoration » E_{-1} nous obtiendrons à partir du vecteur dominant v_0 tous les autres vecteurs. Par conséquent, $U = V_n$, et V_n est un L-module irréductible.

Remarquons que V_0 est un module trivial (de dimension un), alors que V_1 est un module correspondant à la définition naturelle de l'algèbre L : dans la base $\{v_0, v_1\}$ les opérateurs E_{-1}, E_0, E_1 ont pour matrices

$$\begin{Vmatrix} 0 & 0 \\ 1 & 0 \end{Vmatrix}, \quad \begin{Vmatrix} 1 & 0 \\ 0 & -1 \end{Vmatrix}, \quad \begin{Vmatrix} 0 & 1 \\ 0 & 0 \end{Vmatrix}.$$

Le théorème suivant résout le problème que nous avons posé:

THÉORÈME 6.— *Tout L-module V irréductible de dimension $n + 1$ sur $\mathbb{C}$ est isomorphe à V_n.*

DÉMONSTRATION.— D'après le lemme 2, le module V possède un vecteur dominant v_0 de poids λ. Posons

$$v_{-1} = 0 \text{ et } v_m = \frac{1}{m!} E_{-1}^m v_0 = \frac{1}{m!} E_{-1} (\ldots (E_{-1} v) \ldots) \text{ pour } m \geqslant 0.$$

Pour tout $m \geqslant 0$, les formules suivantes sont vraies:

$$\begin{aligned} E_{-1} v_m &= (m + 1)\, v_{m+1}, \\ E_0 v_m &= (\lambda - 2m)\, v_m, \\ E_1 v_m &= (\lambda - m + 1)\, v_{m-1}. \end{aligned} \tag{10'}$$

En effet, pour $m = 0$, les formules (10′) se ramènent à la définition du vecteur dominant v_0 et du vecteur v_1. Puis, raisonnons par récurrence sur m : a) la formule $E_{-1} v_m = (m + 1)\, v_{m+1}$ définit le vecteur v_{m+1} ; b) la formule $E_0 v_m = (\lambda - 2m)\, v_m$ résulte du lemme 1; c) s'il est déjà connu que $E_1 v_{m-1} = (\lambda - m + 2)\, v_{m-2}$, alors en simplifiant par m les deux membres de l'égalité

$$\begin{aligned} mE_1 v_m = E_1 (E_{-1} v_{m-1}) &= \\ &= [E_1 E_{-1}]\, v_{m-1} + E_{-1} (E_1 v_{m-1}) = \\ &= E_0 v_{m-1} + (\lambda - m + 2)\, E_{-1} v_{m-2} = \\ &= \{(\lambda - 2m + 2) + (\lambda - m + 2)(m - 1)\}\, v_{m-1} = \\ &= m (\lambda - m + 1)\, v_{m-1} \end{aligned}$$

on obtient la dernière formule de (10′).

Si les vecteurs $v_0, v_1, \ldots, v_r$ sont différents de zéro pour un r quelconque, alors, étant de poids différents, ils doivent être linéairement indépendants. D'autre part, le module V étant irréductible, le sous-module engendré par le vecteur v_0 coïncide avec V, et comme $\dim V = n + 1$, on a $V = \langle v_0, v_1, \ldots, v_n \rangle$ et $v_{n+1} = v_{n+2} = \ldots$

$\ldots = 0$. En particulier,

$$0 = E_1 v_{n+1} = (\lambda - n)\, v_n = 0 \Rightarrow \lambda = n$$

(on attire l'attention sur une implication bien curieuse : $\dim V < \infty \Rightarrow \lambda \in \mathbb{Z},\ \lambda \geqslant 0$).

En portant la valeur de $\lambda = n$ dans les formules (10′), nous retrouvons au fait, compte tenu des notations adoptées, les formules (9) qui définissent le L-module V_n. Par suite, $V \cong V_n$. ■

EXERCICES

1. L'équation $x^2 + 1 = 0$ combien de solutions admet-elle dans l'algèbre des quaternions $\mathbb{H}$?

2. *Algèbre des quaternions généralisés sur* $\mathbb{Q}$. Montrer que la table de multiplication

	1	e_1	e_2	e_3
1	1	e_1	e_2	e_3
e_1	e_1	n	e_3	ne_2
e_2	e_2	$-e_3$	m	$-me_1$
e_3	e_3	$-ne_2$	me_1	$-nm$

avec $n, m \in \mathbb{Z}$, $nm \neq 0$, définit sur l'espace vectoriel $\mathbb{H}(n, m) = \langle 1, e_1, e_2, e_3\rangle_{\mathbb{Q}}$ de dimension quatre sur $\mathbb{Q}$ une structure d'algèbre associative ayant un élément unité. A cet effet, utiliser la représentation

$$x = x_0 + x_1 e_1 + x_2 e_2 + x_3 e_3 \mapsto A_x = \begin{Vmatrix} x_0 + x_1\sqrt{n} & x_2\sqrt{m} + x_3\sqrt{nm} \\ x_2\sqrt{m} - x_3\sqrt{mn} & x_0 - x_1\sqrt{n}. \end{Vmatrix}.$$

Le déterminant $\det A_x = x_0^2 - x_1^2 n - x_2^2 m + x_3^2 nm = N(x)$ s'appelle norme de l'élément x. Vérifier que, si la condition $x \in \mathbb{H}(n, m)$, $x \neq 0 \Rightarrow N(x) \neq 0$ est satisfaite, l'espace $\mathbb{H}(n, m)$ est une algèbre à division (*algèbre généralisée des quaternions*). En se servant des notions et des résultats de l'exercice 7 du § 2, montrer que dans le cas où $p \equiv \pm 3 \pmod{8}$ est premier, l'algèbre $\mathbb{H}(2, p)$ sera une algèbre à division.

3. Considérer $\mathbb{F}_{2^n}$ comme un espace vectoriel V de dimension n sur $\mathbb{F}_2$. En plus de l'opération d'addition qui provient de $\mathbb{F}_{2^n}$, introduire sur V une opération de multiplication $(x, y) \mapsto x \circ y = \sqrt{xy}$. Ici, $x \mapsto \sqrt{x}$ est un automorphisme sur $\mathbb{F}_{2^n}$, inverse de $x \mapsto x^2$, de sorte que $\sqrt{x+y} = \sqrt{x} + \sqrt{y}$. Montrer que $(V, +, \circ)$ est une algèbre commutative (non associative) sur $\mathbb{F}_2$, qui vérifie les propriétés suivantes : a) V ne contient ni diviseurs de zéro, ni élément unité ; b) l'équation $a \circ x = b$, avec $a \neq 0$, est univoquement résoluble ; c) le groupe des automorphismes $\mathrm{Aut}(V)$ opère transitivement sur $V \setminus \{0\}$.

4. Dans toute algèbre est vérifiée l'identité

$$t(x, y, z) + (t, x, y)z = (tx, y, z) - (t, xy, z) + (t, x, yz).$$

S'en assurer par une vérification directe et montrer que, si dans une algèbre A, ayant un élément unité 1, sur un corps commutatif P, on a pour tous les associateurs l'appartenance $(x, y, z) \in P \cdot 1$, alors A est une algèbre associative.

ANNEXE

FORME RÉDUITE DE JORDAN DES MATRICES

Cet «îlot » d'algèbre linéaire n'est légèrement éclairé ici que pour souligner sa similitude avec le § 5 du chapitre 7, où a été donnée la classification des groupes abéliens finis. Nous n'avons pas cru nécessaire d'insister au chapitre 9, § 3 sur le rôle unificateur que jouent dans cette question les modules sur les anneaux principaux, parce qu'aux différentes catégories de lecteurs il sera peut-être plus commode d'avoir des preuves directes des faits relatifs aux groupes et aux opérateurs linéaires.

1. En cherchant à comprendre comment agit un opérateur linéaire donné $\mathcal{A} : V \to V$, il est naturel de se proposer de trouver dans V une base qui s'accorde au mieux avec cet opérateur. Autrement dit, il faut rechercher dans la classe de matrices semblables $C^{-1}AC$ associées à l'opérateur $\mathcal{A}$ une matrice se présentant sous une forme aussi simple que possible. Pour des raisons bien compréhensibles, ce problème est étroitement lié au corps commutatif de base P sur lequel est défini l'espace vectoriel V. Dans ce qui suit, on suppose que $P = \mathbb{C}$ est le corps des nombres complexes ou un corps algébriquement clos quelconque.

Soient $n = \dim V$ et $\lambda_1, \ldots, \lambda_n$ les racines du polynôme caractéristique

$$f_{\mathcal{A}}(t) = f_A(t) = \det(tE - A) = t^n + a_1 t^{n-1} + \ldots + a_n = \prod_{i=1}^{n} (t - \lambda_i),$$

$$a_1 = -\operatorname{tr} A = -(\lambda_1 + \ldots + \lambda_n),$$

$$a_n = (-1)^n \det A = (-1)^n \lambda_1 \ldots \lambda_n.$$

Les nombres complexes λ_i sont aussi des valeurs propres de l'opérateur linéaire $\mathcal{A}$: les sous-espaces

$$V_{\lambda_i} = \{v \in V \mid \mathcal{A}v = \lambda_i v\}$$

sont différents de zéro, et leurs vecteurs non nuls s'appellent vecteurs propres de l'opérateur $\mathcal{A}$. L'ensemble Spec $(\mathcal{A})$ de toutes les valeurs propres deux à deux distinctes (des racines caractéristiques)

de l'opérateur $\mathcal{A}$ s'appelle *spectre* de $\mathcal{A}$. On définit de manière analogue le *spectre* Spec (A) *de la matrice* A.

Indiquons les faits suivants:

(i) *Les vecteurs propres associés aux valeurs propres différentes sont linéairement indépendants. La somme* $\sum_{\lambda \in \operatorname{Spec}(A)} V^\lambda$ *est directe* (en général, $\sum V^\lambda$ ne coïncide pas avec V).

(ii) *La matrice d'un opérateur linéaire* $\mathcal{A}$ *peut toujours être réduite (au sens de similitude) à une forme triangulaire.*

On peut s'en assurer de façon très simple en raisonnant par récurrence. Il faut prendre un sous-espace à une dimension $\mathcal{A}$-invariant $\langle e_1 \rangle$ $(\mathcal{A} e_1 = \lambda_1 e_1)$, passer à l'espace quotient $\overline{V} = V/\langle e_1 \rangle = \{\overline{v} = v + \langle e_1 \rangle \mid v \in V\}$ de dimension $n - 1$ et à l'opérateur quotient $\overline{\mathcal{A}} : \overline{\mathcal{A}}\,\overline{v} = \overline{\mathcal{A}v}$, choisir dans $\overline{V}$ une base $\overline{e}_2, \ldots, \overline{e}_n$ qui réduit $\overline{A}$ à la forme triangulaire et revenir à l'espace V:

$$A = \begin{Vmatrix} \lambda_1 & & & * \\ & \lambda_2 & & \\ 0 & & \ddots & \\ & & & \lambda_n \end{Vmatrix}. \ ■$$

(iii) (Théorème de Cayley-Hamilton).— *L'opérateur linéaire* $\mathcal{A}$ *et la matrice* A *qui lui correspond (dans toute base) sont annulés par leur polynôme caractéristique.*

Cette dernière assertion étant indépendante du choix de la base, il est commode d'utiliser la propriété (ii). Considérons la suite de sous-espaces $\mathcal{A}$-invariants $V = V_0 \supset V_1 \supset \ldots \supset V_{n-1} \supset 0$, où $V_k = \langle e_1, e_2, \ldots, e_{n-k-1}, e_{n-k} \rangle$. Puisque $(\mathcal{A} - \lambda_{n-k}\mathcal{E})\, e_{n-k} \in V_{k+1}$, on a $(\mathcal{A} - \lambda_{n-k}\,\mathcal{E})\, V_k \subset V_{k+1}$ et donc

$$f_{\mathcal{A}}(\mathcal{A})\, V = \prod_{i=1}^{n} (\mathcal{A} - \lambda_i \mathcal{E})\; V =$$
$$= (\mathcal{A} - \lambda_1 \mathcal{E}) \ldots (\mathcal{A} - \lambda_n \mathcal{E})\, V_0 \subset (\mathcal{A} - \lambda_1 \mathcal{E}) \ldots (\mathcal{A} - \lambda_{n-1} \mathcal{E})\, V_1 \subset$$
$$\subset (\mathcal{A} - \lambda_1 \mathcal{E}) \ldots (\mathcal{A} - \lambda_{n-2} \mathcal{E})\, V_2 \subset \ldots \subset (\mathcal{A} - \lambda_1 \mathcal{E})\, V_{n-1} = 0.$$

Or, $f_{\mathcal{A}}(\mathcal{A})\, V = 0 \Leftrightarrow f_{\mathcal{A}}(\mathcal{A}) = 0$. ■

(iv) *Le polynôme minimal* $h_{\mathcal{A}}(t) = h_A(t)$ *de l'opérateur (polynôme unitaire de degré minimal* $m \leqslant n$, *annulant* $\mathcal{A}$ *et* A*) divise le polynôme caractéristique* $f_{\mathcal{A}}(t)$ *et est divisible par tous les facteurs linéaires* $t - \lambda$, $\lambda \in$ Spec $(\mathcal{A})$.

La division euclidienne $f_{\mathcal{A}}(t) = q(t) \cdot h_{\mathcal{A}}(t) + r(t)$, $\deg r(t) < \deg h_{\mathcal{A}}(t)$, et les propriétés $f_{\mathcal{A}}(\mathcal{A}) = 0 = h_{\mathcal{A}}(\mathcal{A})$ montrent que

$r(\mathcal{A}) = 0$, d'où $r(t) = 0$. Si λ est valeur propre de l'opérateur $\mathcal{A}$, alors $\mathcal{A}v = \lambda v \Rightarrow 0 = h_{\mathcal{A}}(\mathcal{A})v = h_{\mathcal{A}}(\lambda)v \Rightarrow h_{\mathcal{A}}(\lambda) = 0 \Rightarrow (t - \lambda) \mid \times$ $\times\, h_{\mathcal{A}}(t)$.

EXEMPLE.— On dit qu'un opérateur linéaire $\mathcal{A}: V \to V$ est *nilpotent*, si $\mathcal{A}^m = 0$; m est l'*indice de nilpotence* si $\mathcal{A}^{m-1} \neq 0$. *Soit $\mathcal{A}^{m-1}v \neq 0$. Alors, les vecteurs $v, \mathcal{A}v, \ldots, \mathcal{A}^{m-1}v$ sont linéairement indépendants.* En effet, toute dépendance linéaire non triviale est de la forme

$$\mathcal{A}^k v + \alpha_1 \mathcal{A}^{k+1} v + \ldots + \alpha_{m-1-k} \mathcal{A}^{m-1} v = 0, \quad 0 \leqslant k \leqslant m-1.$$

L'application de l'opérateur $\mathcal{A}^{m-1-k}$ aux deux membres de cette égalité nous conduirait à la relation $\mathcal{A}^{m-1}v = 0$ qui contredit le choix de v.

Ainsi, l'indice de nilpotence m de l'opérateur $\mathcal{A}$ est au plus égal à $n =$ $= \dim V$. Soient $m = n$ et $\mathcal{A}^{n-1}v \neq 0$. Introduisons pour les vecteurs de base les désignations suivantes: $v_1 = \mathcal{A}^{n-1}v$, $v_2 = \mathcal{A}^{n-2}v, \ldots, v_{n-1} = \mathcal{A}v$, $v_n = v$. Alors, $\mathcal{A}v_1 = 0$, $\mathcal{A}v_k = v_{k-1}$, $k > 1$, et la matrice de l'opérateur $\mathcal{A}$ dans la base $\{v_1, \ldots, v_n\}$ sera une matrice de Jordan:

$$J_{n,0} = \begin{Vmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{Vmatrix}.$$

Si par exemple, $V = \langle 1, X, X^2, \ldots, X^{n-1} \rangle_{\mathbb{C}}$ est l'espace des polynômes de degré $< n$ sur $\mathbb{C}$, et $\mathcal{A} = \frac{d}{dX}$ est l'opérateur de dérivation, la matrice de cet opérateur dans la base $\{e_i\}$, $e_i = \frac{1}{i!} X^i$, sera justement la matrice $J_{n,0}$.

Plus généralement, nous appellerons *matrice de Jordan* (*supérieure*) de type $m \times m$ (ou d'ordre m), correspondant à la valeur propre λ, une matrice de la forme

$$J_{m,\lambda} = \begin{Vmatrix} \lambda & 1 & 0 & \ldots & 0 & 0 \\ 0 & \lambda & 1 & \ldots & 0 & 0 \\ 0 & 0 & \lambda & \ldots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & \lambda & 1 \\ 0 & 0 & 0 & \ldots & 0 & \lambda \end{Vmatrix}.$$

Remarquons que $J_{m,\lambda} - \lambda E = J_{m,0}$ est une matrice nilpotente: $J_{m,0}^{m-1} \neq 0$, $J_{m,0}^m = 0$.

En particulier, $(t - \lambda)^m$ est le polynôme minimal de la matrice de Jordan $J_{m,\lambda}$, et λ sa valeur propre unique: $\mathrm{Spec}\,(J_{m,\lambda}) = \{\lambda\}$.

Si $u(t)$ est un polynôme arbitraire, on a

$$u(J_{m,\lambda}) = \begin{Vmatrix} u(\lambda) & u'(\lambda)/1! & u''(\lambda)/2! & \dots & u^{(m-1)}(\lambda)/(m-1)! \\ 0 & u(\lambda) & u'(\lambda)/1! & \dots & u^{(m-2)}(\lambda)/(m-2)! \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & u(\lambda) \end{Vmatrix},$$

si bien qu'il est plus facile d'opérer avec $J_{m,\lambda}$ qu'avec des matrices quelconques.

Théorème fondamental.— *Toute matrice carrée A d'ordre n sur un corps commutatif algébriquement clos P (en particulier sur $\mathbb{C}$) est semblable à une somme directe de matrices de Jordan. A savoir, il existe une matrice régulière C telle que*

$$C^{-1}AC = J_{m_1,\lambda_1} \dotplus \dots \dotplus J_{m_s\ \lambda_s} = \begin{Vmatrix} J_{m_1,\lambda_1} & & & 0 \\ & J_{m_2,\lambda_2} & & \\ & 0 & \ddots & \\ & & & J_{m_s,\lambda_s} \end{Vmatrix}$$

(forme réduite de Jordan $J(A)$ de la matrice A). La forme réduite de Jordan est unique à une permutation des matrices de Jordan J_{m_i,λ_i} près.

Comme les polynômes minimaux des matrices semblables coïncident, du théorème fondamental et des remarques faites sur la matrice de Jordan $J_{m,\lambda}$ il résulte que

$$h_A(t) = (t - \lambda_{i_1})^{m_{i_1}} \dots (t - \lambda_{i}\ )^{m_{i_p}},$$

où $\{\lambda_{i_1}, \dots, \lambda_{i_p}\} = \mathrm{Spec}\,(A)$ et m_{j_k} est l'ordre maximal de la matrice de Jordan correspondant à la valeur propre λ_{j_k}.

Il est clair qu'une condition nécessaire et suffisante pour qu'une matrice A soit diagonalisable (c'est-à-dire semblable à la matrice diag $\{\lambda_1, \dots, \lambda_n\}$ est l'absence dans $J(A)$ de matrices de Jordan J_{m_i,λ_i} d'ordre supérieur à 1. On peut donc énoncer un critère bien utile suivant:

Corollaire.— *Une matrice carrée A sur $\mathbb{C}$ est diagonalisable si, et seulement si, son polynôme minimal $h_A(t)$ n'a pas de racines multiples.*

Ce critère est efficace parce que pour calculer $h_A(t)$ il n'est pas besoin de représenter la matrice A sous la forme réduite de Jordan.

La démonstration du théorème fondamental est partagée en trois parties correspondant aux n[os] 2, 3, 4.

2. L'ensemble des vecteurs

$$V(\lambda) = \{v \in V \mid (\mathcal{A} - \lambda\mathcal{E})^k v = 0 \text{ pour un certain } k\}$$

s'appelle *sous-espace de racines* correspondant à la valeur propre $\lambda \in \operatorname{Spec}(\mathcal{A})$.

Une vérification facile à effectuer nous assure que $V(\lambda)$ est en effet un sous-espace. Si, par exemple $u \in V(\lambda)$, $v \in V(\lambda)$ sont tels que $(\mathcal{A} - \lambda\mathcal{E})^s u = 0$, $(\mathcal{A} - \lambda\mathcal{E})^t v = 0$ et $m = \max\{s, t\}$, alors on a

$$(\mathcal{A} - \lambda\mathcal{E})^m (\alpha u + \beta v) = \alpha (\mathcal{A} - \lambda\mathcal{E})^m u + \beta (\mathcal{A} - \lambda\mathcal{E})^m v = 0,$$

d'où $\alpha u + \beta v \in V(\lambda)$, quels que soient $\alpha, \beta \in \mathbb{C}$. Puisque $V(\lambda)$ contient un vecteur propre correspondant *à* λ, on a $V(\lambda) \neq 0$. Puis, $V^\lambda \subset V(\lambda)$; ici, l'égalité peut fort bien ne pas avoir lieu comme le montre l'exemple d'un opérateur nilpotent $\mathcal{A}$ d'indice de nilpotence n, que nous avons considéré plus haut. Dans ce cas, $\lambda = 0$ est l'unique valeur propre, $\dim V^0 = 1$, mais $V(0) = V$.

Puisque $\dim V(\lambda) \leqslant n$ et la restriction de $\mathcal{A} - \lambda\mathcal{E}$ à $V(\lambda)$ est un opérateur nilpotent, on a

$$V(\lambda) = \{v \in V \mid (\mathcal{A} - \lambda\mathcal{E})^n v = 0\}.$$

THÉORÈME 1.— *Soit $\mathcal{A}: V \to V$ un opérateur linéaire de polynôme caractéristique*

$$f_{\mathcal{A}}(t) = \prod_{i=1}^{p} (t - \lambda_i)^{n_i} \quad (\lambda_i \neq \lambda_j \quad \text{pour} \quad i \neq j).$$

Alors, $V = V(\lambda_1) \oplus \ldots \oplus V(\lambda_p)$ est somme directe des sous-espaces de racines $V(\lambda_i)$ dont chacun est invariant par $\mathcal{A}$ et est de dimension $\dim V(\lambda_i) = n_i$. *L'opérateur $\mathcal{A} - \lambda_i\mathcal{E}$, nilpotent sur $V(\lambda_i)$ agit de façon non dégénérée sur le sous-espace*

$$V_i = V(\lambda_i) \oplus \ldots \oplus V(\lambda_{i-1}) \oplus V(\lambda_{i+1}) \oplus \ldots \oplus V(\lambda_p).$$

Enfin, λ_i est l'unique valeur propre de l'opérateur $\mathcal{A}|_{V(\lambda_i)}$.

DÉMONSTRATION.— Aucun des facteurs premiers $t - \lambda_k$ ne peut être simultanément un diviseur de tous les polynômes

$$f_i(t) = \prod_{j \neq i} (t - \lambda_j)^{n_j}, \quad i = 1, 2, \ldots, p,$$

et, de ce fait, P.G.C.D. $(f_1(t), \ldots, f_p(t)) = 1$. Par suite, il existe des polynômes $g_1(t), \ldots, g_p(t) \in \mathbb{C}[t]$, pour lesquels

$$\sum_{i=1}^{p} f_i(t)\, g_i(t) = 1. \tag{1}$$

Les sous-espaces

$$W_i = f_i(\mathcal{A})\, g_i(\mathcal{A})\, V = \{f_i(\mathcal{A})\, g_i(\mathcal{A})\, v \mid v \in V\}, \quad 1 \leqslant i \leqslant p,$$

sont invariants par rapport à $\mathcal{A}$:

$$\mathcal{A} W_i = f_i(\mathcal{A})\, g_i(\mathcal{A})\, \mathcal{A} V \subset f_i(\mathcal{A})\, g_i(\mathcal{A})\, V = W_i.$$

De plus

$$(\mathcal{A} - \lambda_i \mathcal{E})^{n_i} W_i = f_{\mathcal{A}}(\mathcal{A})\, g_i(\mathcal{A})\, V = 0$$

(puisque $f_{\mathcal{A}}(\mathcal{A}) = 0$; voir (iii)), si bien que

$$W_i \subset V(\lambda_i). \tag{2}$$

La relation (1) mise sous la forme

$$\mathcal{E} = \sum_{i=1}^{p} f_i(\mathcal{A})\, g_i(\mathcal{A}),$$

nous donne la décomposition

$$V = \sum_{i=1}^{p} W_i \tag{3}$$

et à plus forte raison (par suite de l'inclusion (2)):

$$V = \sum_{i=1}^{p} V(\lambda_i).$$

Supposons que $v \in V(\lambda_i) \cap V_i$, où, de même que dans l'énoncé du théorème, $V_i = \sum_{j \neq i} V(\lambda_j)$. Alors, $(\mathcal{A} - \lambda_i \mathcal{E})^n v = 0$ et, puisque $v = \sum_{j \neq i} v_j$ et $(\mathcal{A} - \lambda_j \mathcal{E})^n v_j = 0$, on a $\{\prod_{j \neq i} (\mathcal{A} - \lambda_j \mathcal{E})^n\} v = 0$. Or, du fait que les polynômes $(t - \lambda_i)^n$, $c(t) = \prod_{j \neq i} (t - \lambda_j)^n$ sont premiers entre eux, il résulte l'existence de $a(t)$, $b(t)$, tels que

$$a(t)\,(t - \lambda_i)^n + b(t)\, c(t) = 1.$$

Il vient:

$$v = a(\mathcal{A})\,(\mathcal{A} - \lambda_i)^n v + b(\mathcal{A})\,\{\prod_{j \neq i} (\mathcal{A} - \lambda_j \mathcal{E})^n\}\, v = 0,$$

c'est-à-dire les espaces $V(\lambda_i)$ et V_i sont disjoints. Ceci signifie que nous avons une décomposition

$$V = V(\lambda_1) \oplus \ldots \oplus V(\lambda_p) \tag{4}$$

en somme directe de sous-espaces $\mathcal{A}$-invariants.

De l'inclusion (2) et des décompositions (3) et (4), il ressort immédiatement que $W_i = V(\lambda_i)$. Ainsi, nous avons obtenu pour $V(\lambda_i)$

une expression efficace

$$V(\lambda_i) = f_i(\mathcal{A})\, g_i(\mathcal{A})\, V,$$

où $f_i(t)$, $g_i(t)$ sont les polynômes intervenant dans l'identité (1). En particulier,

$$(\mathcal{A} - \lambda_i)^{n_i} V(\lambda_i) = 0.$$

L'opérateur $\mathcal{A}$ a pour polynôme minimal sur $V(\lambda_i)$ un diviseur du polynôme $(t - \lambda_i)^{n_i}$. Il en résulte premièrement que λ_i est l'unique valeur propre de l'opérateur $\mathcal{A}\,|_{V(\lambda_i)}$. Deuxièmement, dans la base qui est réunion des bases des espaces $V(\lambda_i)$, l'opérateur $\mathcal{A}$ est de matrice

$$A = \begin{Vmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_p \end{Vmatrix},$$

où A_i est une matrice d'ordre $n'_i = \dim V(\lambda_i)$, ayant l'unique valeur propre λ_i et un polynôme caractéristique $f_{A_i}(t) = (t - \lambda_i)^{n'_i}$, $n'_i \leqslant n_i$. Puisque $f_A(t) = \prod_{i=1}^{p} f_{A_i}(t)$ on a $n = n'_1 + \ldots + n'_p$ et $n'_i = n_i$.

Il nous reste à démontrer que la restriction $(\mathcal{A} - \lambda_i \mathcal{E})|_{V_i}$ est non dégénérée. Or, c'est facile à comprendre, car dans le cas contraire on aurait $\{\operatorname{Ker}(\mathcal{A} - \lambda_i \mathcal{E})\} \cap V_i \neq 0$ et $\mathcal{A}v - \lambda_i v = 0$ pour un $0 \neq v \in V_i$. Pourtant, sur V_i, le polynôme caractéristique de $\mathcal{A}$ est $f_i(t) = \prod_{j \neq i} (t - \lambda_j)^{n_j}$ et λ_i ne peut pas être la valeur propre.

3. Le théorème 1 que nous venons de démontrer ramène le problème du choix de la matrice la plus simple pour l'opérateur linéaire $\mathcal{A} : V \to V$ au cas où $\mathcal{A}$ possède l'unique valeur propre λ, et $(\mathcal{A} - \lambda \mathcal{E})^m = 0$, $m \leqslant \dim V$. En posant $\mathcal{B} = \mathcal{A} - \lambda \mathcal{E}$, nous obtiendrons un opérateur nilpotent d'indice de nilpotence m avec une matrice nilpotente B.

THÉORÈME 2.— *La forme réduite de Jordan $J(B)$ d'une matrice nilpotente B existe* (le corps commutatif de base P est arbitraire).

DÉMONSTRATION.— Il nous faut montrer que l'espace vectoriel V sur lequel agit un opérateur nilpotent $\mathcal{B}$ de matrice B, se décompose en une somme directe de *sous-espaces* dits *cycliques* $P[\mathcal{B}]\, v_i = \langle v_i, \mathcal{B}v_i, \ldots, \mathcal{B}^{m_i - 1} v_i \rangle$, avec $\mathcal{B}^{m_i} v_i = 0$. Nous voulons raisonner par récurrence sur la dimension de l'espace. Supposons l'assertion du théorème démontrée pour tous les couples $(V', \mathcal{B}')$, où $\dim V' < \dim V$ et $\mathcal{B}'$ est un opérateur nilpotent sur V'.

Soient $\mathscr{B}^m = 0$, $\mathscr{B}^{m-1}u \neq 0$. Introduisons le sous-espace cyclique $U = \langle u, \mathscr{B}u, \ldots, \mathscr{B}^{m-1} u\rangle$ et l'espace quotient $\overline{V} = V/U$ sur lequel nous ferons agir l'opérateur quotient $\overline{\mathscr{B}}$: $\overline{\mathscr{B}}\overline{v} = \overline{\mathscr{B}v}$. Ici $\overline{v} = v + U$ est une classe ayant v pour représentant. Puisque $\overline{\mathscr{B}}^m\overline{v} = \overline{\mathscr{B}^m v} = 0$, $\overline{\mathscr{B}}$ est un opérateur nilpotent d'indice de nilpotence $\overline{m} \leqslant m$. Autrement dit, $\mathscr{B}^{\overline{m}-1}V \not\subset U$, $\mathscr{B}^{\overline{m}}V \subseteq U$.

Puisque $\dim \overline{V} < \dim V$, on a, par hypothèse de récurrence,

$$\overline{V} = \overline{U}_1 \oplus \ldots \oplus \overline{U}_{s-1}, \quad \overline{U}_i = P\,[\overline{\mathscr{B}}]\,\overline{u}_i.$$

On obtient pour V une décomposition

$$V = U_1 \oplus \ldots \oplus U_{s-1} \oplus U, \tag{5}$$

où

$$U_i = \langle u_i, \mathscr{B}u_i, \ldots, \mathscr{B}^{m_i-1}u_i\rangle, \quad \mathscr{B}^{m_i}u_i \in U, \quad m_i \leqslant \overline{m} \leqslant m.$$

Les sous-espaces U_i ne sont pas $\mathscr{B}$-invariants, car en général $\mathscr{B}^{m_i}u_i \neq 0$.

Pour la commodité des notations, i étant fixe, posons $w = u_i$, $l = m_i$, $W = U_i = \langle w, \mathscr{B}w, \ldots, \mathscr{B}^{l-1}w\rangle$. Par hypothèse

$$\mathscr{B}^l w = \alpha_k \mathscr{B}^k u + \alpha_{k+1}\mathscr{B}^{k+1}u + \ldots + \alpha_{m-1}\mathscr{B}^{m-1}u, \quad \alpha_k \neq 0$$

(si tous les $\alpha_j = 0$, on n'a rien à faire). En appliquant à la dernière relation l'opérateur $\mathscr{B}^{m-1-k}$, nous obtiendrons $\mathscr{B}^{m-1-k+l}w = \alpha_k\mathscr{B}^{m-1}u \neq 0$. Comme $\mathscr{B}^m = 0$, ceci ne peut avoir lieu que pour $l \leqslant k \leqslant m-1$. En posant

$$v = w - \alpha_k\mathscr{B}^{k-l}u - \alpha_{k+1}\mathscr{B}^{k-l+1}u - \ldots - \alpha_{m-1}\mathscr{B}^{m-1-l}u,$$

nous constatons que $\mathscr{B}^{l-1}v = \mathscr{B}^{l-1}w + u' \neq 0$, mais

$$\mathscr{B}^l v = \mathscr{B}^l w - \alpha_k\mathscr{B}^k u - \ldots - \alpha_{m-1}\mathscr{B}^{m-1}u = 0.$$

L'espace cyclique $\langle v, \mathscr{B}v, \ldots, \mathscr{B}^{l-1}v\rangle$, avec $\mathscr{B}^l v = 0$, engendre avec U le sous-espace $U_i \oplus U$.

Ces raisonnements étant valables pour tout i, $1 \leqslant i \leqslant s-1$, nous pouvons remplacer dans la décomposition (5) chaque sous-espace U_i par $V_i = \langle v_i, \mathscr{B}v_i, \ldots, \mathscr{B}^{m_i-1}v_i\rangle$, $\mathscr{B}^{m_i}v_i = 0$. En posant encore $v_s = u$, $m_s = m$, $V_s = U$, nous obtenons la décomposition

$$V = V_1 \oplus \ldots \oplus V_s$$

qui vérifie toutes les propriétés requises.

4. En abordant la démonstration de l'unicité, indiquons au passage une *règle pratique permettant de mettre une matrice arbitraire A d'ordre n sous la forme réduite de Jordan.*

A cet effet, il faut apprendre à déterminer le nombre $N(m, \lambda)$ de matrices de Jordan $J_{m,\lambda}$ d'ordre m correspondant à la valeur propre λ de la matrice A. A la matrice A faisons correspondre de façon habituelle un opérateur $\mathscr{A}$ agissant sur un espace vectoriel V de dimension n, et décomposons V en une somme directe

$$V = V(\lambda) \oplus V', \tag{6}$$

où

$$V(\lambda) = \bigoplus_{j=1}^{s} \langle v_j, (\mathscr{A} - \lambda\mathscr{E}) v_j, \ldots, (\mathscr{A} - \lambda\mathscr{E})^{m_j-1} v_j \rangle, \quad V' = \sum_{\lambda' \neq \lambda} V(\lambda').$$

Calculons le rang $r_t = \text{rang}\,(A - \lambda E)^t$ de la matrice $(A - \lambda E)^t$ ou, ce qui revient au même, la dimension de l'espace $(\mathscr{A} - \lambda\mathscr{E})^t V$. Cette dimension est évidemment indépendante du choix de la base dans V. Chacun des espaces intervenant dans la décomposition (6) étant invariant par rapport à $(\mathscr{A} - \lambda\mathscr{E})^t$, on a

$$\dim (\mathscr{A} - \lambda\mathscr{E})^t V = \sum \dim (\mathscr{A} - \lambda\mathscr{E})^t \mathbb{C}[\mathscr{A}] v_j + \dim (\mathscr{A} - \lambda\mathscr{E})^t V'.$$

Pour plus de détermination, supposons $m_1 \leqslant m_2 \leqslant \ldots \leqslant m_s$. Si $m_j \leqslant t$, alors $(\mathscr{A} - \lambda\mathscr{E})^t \mathbb{C}[\mathscr{A}] v_j = 0$. Pour $m_j > t$, on a

$$(\mathscr{A} - \lambda\mathscr{E})^t \mathbb{C}[\mathscr{A}] v_j =$$
$$= \langle (\mathscr{A} - \lambda\mathscr{E})^t v_j, (\mathscr{A} - \lambda\mathscr{E})^{t+1} v_j, \ldots, (\mathscr{A} - \lambda\mathscr{E})^{m_j-1} v_j \rangle,$$

si bien que $\dim (\mathscr{A} - \lambda\mathscr{E})^t \mathbb{C}[\mathscr{A}] v_j = m_j - t$. L'opérateur $\mathscr{A} - \lambda\mathscr{E}$ étant non dégénéré sur V' (théorème 1), $\dim (\mathscr{A} - \lambda\mathscr{E})^t V' = \dim V'$. Il vient

$$r_t = \sum_{m_j > t} (m_j - t) + \dim V',$$

d'où

$$r_t - r_{t+1} = \sum_{m_j > t} (m_j - t) - \sum_{m_j > t+1} (m_j - t - 1) =$$
$$= \sum_{m_j > t} (m_j - t) - \sum_{m_j > t+1} (m_j - t) + \sum_{m_j > t+1} 1 =$$
$$= \sum_{m_j = t+1} 1 + \sum_{m_j > t+1} 1 = N(t+1, \lambda) + N(t+2, \lambda) + \ldots$$

Par suite $r_{m-1} - r_m - (r_m - r_{m+1}) = \{N(m, \lambda) + N(m+1, \lambda) + \ldots\} - \{N(m+1, \lambda) + N(m+2, \lambda) + \ldots\} = N(m, \lambda)$, et, en définitive, on obtient la formule

$$\begin{gathered} N(m, \lambda) = r_{m-1} - 2r_m + r_{m+1}, \\ m \geqslant 1, \quad r_t = \text{rang}\,(A - \lambda E)^t, \quad r_0 = n. \end{gathered} \tag{7}$$

Remarquons que r_t est un invariant de la matrice A (c'est-à-dire un nombre défini par la classe de similitude de la matrice A). Ceci

signifie que *la formule* (7) *établit aussi l'unicité de la forme réduite de Jordan* $J(A)$.

Jusqu'ici rien n'a été dit de la matrice C qui assure la similitude

$$J(A) = C^{-1}AC.$$

Mais maintenant, nous connaissons les matrices A et $J(A)$, si bien que $C = (c_{ij})$ peut être définie à partir du système homogène d'équations linéaires

$$CJ(A) - AC = 0$$

d'ordre n^2. Soit $C_1, \ldots, C_r$ son système fondamental de solutions. En général, toutes les C_i ne sont pas des matrices régulières, mais puisque la forme réduite de Jordan $J(A)$ existe, alors $\det(t_1C_1 + \ldots + t_rC_r) \neq 0$ à coefficients indéterminés $t_1, \ldots, t_r$ et on peut choisir $\alpha_1, \ldots, \alpha_r \in \mathbb{C}$ tels que $\det(\alpha_1C_1 + \ldots + \alpha_rC_r) \neq \neq 0$. Alors $C = \alpha_1C_1 + \ldots + \alpha_rC_r$ est la matrice cherchée. Bien entendu, la matrice C est loin d'être définie de manière unique, même lorsqu'elle est normée par $\det C = 1$.

INDEX ALPHABÉTIQUE

E

F

G

H

I

J